

# Resources in Quantum Imaging, Detection and Estimation



Sammy Ragy

School of Mathematical Sciences

University of Nottingham

A thesis submitted for the degree of

*Doctor of Philosophy (PhD)*

April, 2015



## **Abstract**

The research included in this thesis comes in two main bodies. In the first, the focus is on intensity interferometric schemes, and I attempt to identify the types of correlations dominant in their operation. This starts with the, now rather historical, Hanbury Brown and Twiss setup from the 1950s and progresses to more recent interests such as ghost imaging and a variant of ‘quantum illumination’, which is a quantum-enhanced detection scheme. These schemes are considered in the continuous variable regime, with Gaussian states in particular. Intensity interferometry has been the cause of a number of disputes between quantum opticians over the past 60 years and I weigh in on the arguments using relatively recent techniques from quantum information theory.

In the second half, the focus turns away from the optical imaging and detection schemes, and onto quantum estimation – multiparameter quantum estimation to be precise. This is an intriguing area of study where one has to carefully juggle tradeoffs in choosing both the optimal measurement and optimal state for performing an estimation in two or more parameters. I lay out a framework for circumventing some of the difficulties involved in this and apply it to several physical examples, revealing some interesting and at times counterintuitive features of multiparameter estimation.

### **Declaration of originality**

I declare that all work contained herein is my own. Where it is based on published papers, I am the primary author and any points in the text based on significant collaborative work are disclosed as they arise.

## Acknowledgements

I am glad to be able to say that I have learned during these three (and a bit) years of my PhD. I have learned academic things, I have learned about the politics and bureaucratic necessities of academia and I have learned about things outside the academic world entirely. And though there has naturally been some turbulence along the way, the experience has been a positive one, so I have to thank those who have helped make it that way!

I would like to start by thanking my family, who have always been encouraging and reacted with great enthusiasm to any news of publications or other successes, even though I have not yet succeeded at explaining to them quite what it is I do.

The last year and a half of my PhD have been especially valuable, and this is in no small part due to the growth of the quantum information groups at Nottingham (with a steady flux of visitors too). I thus would like to thank the group members, temporary, permanent, or itinerant, who have helped make this possible through group meetings and informal discussions. In as random order as I am capable of, these are Kasia Macieszczak, Sara Di Martino, Mehdi Ahmadi, Giannis Kogias, Luis Correa, Rebecca Schmidt, Jukka Kiukas, Matthew Levitt, Soojoon Lee, Ksenia Samburskaya, Nico Friis, Ivette Fuentes, Jacopo Settimo, Jefferson Filgueiras, Isabela Almeida, Anirudh Acharya, Merlijn van Horssen, Dominik Šafránek, Pietro Liuzzo Scorpo, Tupac Ibarra, Jason Doukas, Antony Lee, Catalin Cătană, Mădălin Guță, Davide Girolami and the diabolical duo Marco Cianciaruso and Tom Bromley. I would like also to thank my officemates and other friends, although I suspect they (on average) were more of an impediment to work than an aid to it.

I would like to thank Rafał Demkowicz-Dobrzański, who I visited in Warsaw to do work on quantum estimation, and who was patient despite delays to that work due to other obligations. His students, Jan Kolodynski and Marcin Jarzyna also provided valuable insights. On this work, I should also extend additional thanks to Mădălin Guță and Katarzyna Macieszczak who helped clarify some concepts for me.

Lastly, I would like to extend special thanks to my supervisor, Gerardo Adesso. Not only has he been always highly supportive, but he has also gone above and beyond to ensure that no important opportunity ever go unseen, whether it be about a conference, a postdoc position or an interesting talk. In addition, I must thank him for his hospitality, with some great pizza evenings at his house (for which I must also thank Samanta and Giovanni!). Although I was one of his first PhD students, I certainly never felt that this had any negative effects at all and I'm grateful for that.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgement</b>	<b>ii</b>
<b>Publications</b>	<b>1</b>
<b>List of Figures</b>	<b>2</b>
<b>Part I</b>	<b>9</b>
<b>1 Introduction</b>	<b>10</b>
<b>2 Toolbox</b>	<b>13</b>
2.1 Quantum States and Observables . . . . .	13
2.1.1 Mixedness . . . . .	15
2.1.2 Evolution . . . . .	17
2.1.3 Composite Systems . . . . .	18
2.1.4 Purification and Dilation . . . . .	21
2.1.5 Generalised Measurement . . . . .	23
2.1.6 Qubits . . . . .	24
2.2 Correlations . . . . .	26
2.2.1 Entropy . . . . .	26
2.2.2 Entanglement . . . . .	28
2.2.3 Mutual Information . . . . .	28
2.2.4 Classical and quantum correlations . . . . .	29
2.2.5 Other entropies . . . . .	32

<b>3</b>	<b>Continuous Variables</b>	<b>34</b>
3.1	Quantum Optics . . . . .	35
3.2	Quasiprobability . . . . .	37
3.2.1	Phase Space . . . . .	40
3.3	Gaussian States . . . . .	42
3.3.1	Operations . . . . .	45
3.3.2	Single and two-mode operations . . . . .	49
3.3.3	Partial tracing and loss . . . . .	51
3.3.4	Gaussian Measurements . . . . .	52
3.4	Entropy and Correlations . . . . .	53
3.4.1	EPR Correlations . . . . .	53
3.4.2	Entropy . . . . .	54
3.4.3	Mutual Information . . . . .	55
3.4.4	Classical and quantum correlations . . . . .	56
<b>Part II</b>		<b>58</b>
<b>4</b>	<b>Intensity Interferometry</b>	<b>59</b>
4.1	Hanbury Brown and Twiss . . . . .	60
4.1.1	Covariance Matrix . . . . .	61
4.1.2	Results . . . . .	65
4.2	Ghost Imaging . . . . .	67
4.2.1	History . . . . .	69
4.2.1.1	Speckle or non-local correlations? . . . . .	71
4.2.2	Correlations in ghost imaging . . . . .	73
4.2.2.1	Lensed Ghost Imaging . . . . .	74
4.2.2.2	Coarse-grained correlations . . . . .	76
4.2.2.3	Results . . . . .	81
4.2.2.4	Discussion . . . . .	85
4.3	Quantum Illumination . . . . .	86
4.3.1	Setup and analysis . . . . .	89
4.3.2	Theoretical results . . . . .	90
4.3.3	Comparisons with experiment . . . . .	93
4.4	Conclusion . . . . .	96

## CONTENTS

---

<b>Part III</b>	<b>99</b>
<b>5 Multiparameter Metrology</b>	<b>99</b>
5.1 Background . . . . .	99
5.1.1 Cramér-Rao bounds . . . . .	102
5.1.2 The Heisenberg Limit . . . . .	104
5.2 Simplifying Multiparameter Estimation . . . . .	107
5.2.1 Holevo Cramér-Rao bound . . . . .	108
5.2.2 Diagonal quantum Fisher information . . . . .	111
5.3 Application . . . . .	113
5.3.1 Preliminary Observations . . . . .	113
5.3.2 Estimation of two unitaries on qubits . . . . .	114
5.3.3 Estimation of phase and dephasing . . . . .	122
5.3.4 Interferometry with Symmetric Loss . . . . .	129
5.3.5 Summary . . . . .	132
<b>6 Conclusions</b>	<b>134</b>
<b>A Gaussian Moment Factoring</b>	<b>137</b>
<b>References</b>	<b>139</b>



# Publications

This thesis contains work from the following publications:

1. Nature of light correlations in ghost imaging, Ragy, S. and Adesso, G., *Scientific Reports* **2** 651 (2012) [1]
2. Unveiling the Hanbury Brown and Twiss effect through Renyi entropy correlations, Ragy, S. and Adesso, G., *Physica Scripta*, **T153** 014052 (2013) [2]
3. Quantifying the source of enhancement in experimental quantum illumination, Ragy, S. et al., *Journal of the Optical Society of America B* **31** 2045 (2014)[3]
4. Continuous variable quantum information: Gaussian states and beyond, Adesso, G., Ragy, S. and Lee, A. R., *Open Systems and Information Dynamics* **21** 1440001 (2014) [4]

In the course of this PhD, the author has also co-authored the following papers which are not included in the thesis:

5. Continuous variable methods in relativistic quantum information: Characterisation of quantum and classical correlations of scalar field modes in noninertial frames, Adesso, G., Ragy, S. and Girolami D., *Classical and Quantum Gravity* **29** 224002 (2012) [5]
6. Continuous variable versus hybrid schemes for quantum teleportation of Gaussian states, Kogias, I., Ragy, S. and Adesso G., *Physical Review A* **89** 052324 (2014) [6]
7. Quantification of Gaussian quantum steering, Kogias, I., Lee, A. R., Ragy, S. and Adesso, G., *Physical Review Letters* **114** 060403 (2015) [7]

# List of Figures

- 2.1 **The Bloch sphere.** For a pure state, where the computational basis is taken to be the eigenbasis of the  $\sigma_z$  Pauli matrix, the two-angle parametrisation is represented as in the diagram. Alternatively, in the vector formalism, the origin is taken to be the centre of the Bloch sphere, and the  $x$ ,  $y$  and  $z$  components are given by the  $\vec{r}$  vector. . . . . 25
- 2.2 **Geometric Quantum Correlations.** An intuitive notion of quantum correlations is gained from classifying states as in the diagram. For some state  $\hat{\rho}$ , we measure the distance to the nearest point in the desired set (QC for discord, for example). The geometric picture above is not representative of the actual shapes or size of the sets in any reasonable sense, but purely for demonstrative purposes [8]. . . . . 32
- 3.1 **Cross-section of the Wigner function for a general pure Gaussian state**  $|\psi_{\alpha,s,\theta}\rangle = \hat{D}(\alpha)\hat{S}(se^{i\theta})|0\rangle$  of a single mode  $k$ , characterised by a complex displacement vector  $\alpha$ , a real squeezing degree  $s$  and a squeezing phase  $\theta$ . The Wigner function is given explicitly by the Gaussian form (3.13b). Figure from [4]. . . . . 45
- 4.1 **(a) Mach-Zehnder and (b) intensity interferometer.** The solid rectangles represent mirrors, and the empty ones represent balanced beam splitters. Despite their depiction, the optical paths in each arm of the detectors need not be the same length. The primary difference between the two is that the Mach-Zehnder interferometer measures the intensity after superposing  $\hat{E}_1$  and  $\hat{E}_2$ , whereas the intensity interferometer measures the intensities separately and then correlates the outputs. 59

4.2	<b>Diagrammatic representation of a Hanbury Brown Twiss setup.</b>	
	Light from a large source propagates into the far field whereupon it is split on a balanced beam-splitter. When one of the detectors is scanned in the transverse plane, the normalised intensity correlations scale as $\frac{2J(\frac{k_0 r}{z} \mathbf{x}-\mathbf{x}' )}{\frac{k_0 r}{z} \mathbf{x}-\mathbf{x}' }$ , where $J$ is a Bessel function of the first kind and first order and $\mathbf{x} - \mathbf{x}'$ is the difference in the transverse positions of the detectors in each arm. Here $r$ refers to the radius of the disc and $z$ the distance of the detectors from disk . . . . .	62
4.3	<b>Plots of the normalised total (<math>\mathcal{T}</math>, solid black line), classical (<math>\mathcal{J}</math>, dashed red line), and quantum (<math>\mathcal{D}</math>, dotted blue line) correlations as a function of <math>x</math> for (a) <math>\bar{n} = 10</math> and (b) <math>\bar{n} = 0.01</math>; the thin green line represents the normalised intensity correlations, which practically coincides with the mutual information in (b). Panel (c) depicts the correlations as a function of <math>\bar{n}</math> with fixed <math>x = 0</math>; the inset shows the same correlations normalised by the mutual information, to better highlight the quantum and classical contributions to the total correlations. In all the plots <math>k_0 r/z = 1000</math>. . . . .</b>	65
4.4	<b>Diagrammatic representation of a ghost imaging scheme . . . .</b>	68
4.5	<b>Completely classical ghost imaging analogue.</b> A double-ended cannon launches projectiles at opposing detectors. In front of a bucket detector lies an object with some gaps through which the projectile may pass. The orientation of the cannon is not directly observable, but the trajectory of both projectiles can be deduced from the position of impact on the spatially resolving detector. When the bucket detector clicks, it can be deduced that the projectile passed through one of the gaps. After a number of launches, the profile of the object can be built up in this way.	70
4.6	<b>An ‘<math>M</math>-splitter’</b> consisting of $M - 1$ beam splitters where the $k$ th beam splitter from the left has a transmissivity of $\sqrt{\frac{k}{k+1}}$ , thus producing an equal combination of all modes. . . . .	77

## LIST OF FIGURES

---

- 4.7 **Conceptual diagrams for coarse-grained correlations**, indicating the areas we are considering on the CCD and bucket planes and a superposition of both. For part (a) we are considering exactly matched areas on either plane, for part (b) we are considering entirely disjoint areas. Part (c) has two areas of equal magnitude, half overlapping. The correlations between these are  $\frac{1}{2}$  those for an individual speckle within the overlapping area. Part (d) considers one area  $\frac{1}{4}$  the size of the other, completely enclosed by it on the superposed picture. This case notably, is most like the comparison of pixel and bucket detector. The effective modes also have  $\frac{1}{2}$  the cross-correlations between them as for any individual speckles within the overlap area. . . . . 80
- 4.8 **Speckle-speckle correlations in thermal-light lensed ghost imaging**, plotted as a function of the illumination  $I$  and the speckle-count  $M$ . In (b) detail at  $M=1$  is shown. Notice the intersection between classical and quantum correlations at  $M = I$ . Classical ones dominate for  $I > M$ , while quantum ones are relevant for  $I < M$ . . . . . 81
- 4.9 **The decomposition of coarse-grained correlations** at  $M = 1$  and  $R = 100$ . Looking at (a) we can see that the total correlations are bounded, and at high illuminations they are mostly classical. Zooming in to the low-illumination regime near the origin, though, we see in (b) that there is an interval where quantum correlations still exceed classical ones. . . . . 82
- 4.10 **Panels [(a)-(b)]: Log-linear plots** of (a) Signal-to-noise ratio (SNR) and (b) total correlations  $\tilde{\mathcal{T}}$  for increasing illumination  $I$ . In panel (c) we display the tiny region filled by SNR versus  $\tilde{\mathcal{T}}$  with varying  $M \in [1, 1000]$  and  $I \in [0, 1000]$ ; the quasi-linear region is essentially invariant upon variations of  $R$  as well. . . . . 83

<p>4.11 <b>(a) Coarse-grained analysis for ghost imaging using entangled light</b> produced by SPDC (with pixel count <math>R = 100</math>): Plot of SNR versus total correlations for <math>M = 1</math> and <math>I \in [0, 1000]</math>; compare with the corresponding thermal-light case, Figure 4.10(c). Panel (b) depicts a comparison between total coarse-grained correlations <math>\tilde{\mathcal{T}}</math> for the cases of thermal and entangled light sources (with <math>R = 100</math>) as a function of the illumination <math>I</math>; it is shown that they share a common limit in the regime of high illumination [see Equation (4.13)], although in this setup SPDC-entangled light is always more correlated than thermal-light. All the correlations are normalised by <math>\sqrt{R/2}</math>. . . . .</p>	84
<p>4.12 <b>A schematic diagram of a quantum illumination scheme.</b> Here the object under consideration is a beam splitter. A portion of the light produced by the correlated source will be reflected by the beam splitter and the challenge is to detect this beneath the dominant thermal noise, thereby discriminating the presence of the object. . . . .</p>	88
<p>4.13 <b>A theoretical plot of the ratios <math>R_{\text{SNR}}</math> (blue solid) and <math>R_{\text{MI}}</math> (red dashed) for quantum illumination</b> with parameters set at realistic experimental values of <math>\langle N_r \rangle = 4000</math>, <math>M = 90000</math>, <math>M_\beta = 50</math>, <math>\eta = 0.38</math> and <math>\eta_\beta = 0.5</math>. The asymptotic limit for <math>\langle N_\beta \rangle \rightarrow \infty</math> is given by Equation (4.19). . . . .</p>	91
<p>4.14 (a)–(d) <b>Experimental results and theoretical expectations for the SNRs (blue solid) and effective MIs (red dashed) of the quantum illumination demonstration of Lopaeva <i>et al.</i> [9] using TWB and THB light, plotted versus the bath photon number <math>\langle N_\beta \rangle</math>.</b> The values of the other experimental parameters are fixed at <math>M = 90000</math>, <math>M_\beta = 1300</math>, <math>\eta = 0.38</math>, <math>\eta_\beta = 0.5</math>, <math>\langle N_{\text{TWB}} \rangle = 4232</math>, <math>\langle N_{\text{THB}} \rangle = 3278</math>, for all the plots in this figure. . . . .</p>	94

## LIST OF FIGURES

---

4.15	(a) <b>Ratios <math>\tilde{R}</math> between SNR and MI</b> for TWB (magenta empty squares) and THB (green filled triangles) sources, obtained from the measured data; the dashed gray curve represents the theoretical prediction, given by the mean between $\tilde{R}_{\text{TWB}}$ and $\tilde{R}_{\text{THB}}$ . (b) Confidence intervals (shadings) inferred from the experimental data, along with theoretical predictions (lines), for the ratios $R_{\text{SNR}}$ (blue solid) and $R_{\text{MI}}$ (red dashed). . . . .	95
5.1	<b>An abstract picture of a metrology scheme.</b> A probe state, $\hat{\rho}_0$ , which may consist of many separable or entangled elements, is subjected to some process determined by the parameters $\theta$ . The resultant state, $\hat{\rho}_\theta$ , is measured with POVM $\{\Pi_x\}$ and the data output is processed to produce an estimate. . . . .	102
5.2	<b>A Mach-Zehnder interferometer for the estimation of phase.</b> The input, $\hat{\rho}$ , may be chosen to be either $N$ separable single photon states, or an entangled NOON state. The output is then either $\frac{1}{\sqrt{2}}( 01\rangle + e^{i\phi} 10\rangle)^{\otimes N}$ or $\frac{1}{\sqrt{2}}( 0N\rangle + e^{iN\phi} N0\rangle)$ respectively. In the former case, we can view this as $N$ independent trials. . . . .	105
5.3	<b>A simple example of an estimation process in which we expect to have non-diagonal Fisher information terms present.</b> Take a source which emits two different particles, one long-lived $\alpha$ and the other a short-lived $\gamma$ which quickly decays into long-lived $\alpha$ and $\beta$ . Take the emissions as normally-distributed random processes dependent on some rates. We cannot directly observe these rates, but must deduce them from some counting statistics. Let's assume that $\gamma$ is too short lived to directly observe, but $\alpha$ and $\beta$ are countable, thus the rate of the $\alpha$ -only production process is only deducible by subtracting the $\beta$ count from the $\alpha$ count. This means the uncertainty in the $\beta$ count will propagate into the estimate of the $\alpha$ -only process, and we can expect a covariance matrix of $\begin{pmatrix} \sigma_\beta^2 & -\sigma_\beta^2 \\ -\sigma_\beta^2 & \sigma_\beta^2 + \sigma_\alpha^2 \end{pmatrix}$ , where $\sigma_{\alpha,\beta}^2$ are the respective variances of the $\gamma$ and $\alpha$ -only processes. Working backwards, this of course also entails non-diagonal Fisher information. . . . .	112

5.4	<b>Sequential waveplates.</b> The first plate has its optical axis aligned vertically, and the second at $\gamma$ degrees to the vertical. Their thicknesses are $t_1$ and $t_2$ respectively. . . . .	116
5.5	<b>A plot of the states which allow saturation of the Cramér-Rao bound.</b> The turquoise (vertical) plane contains the $\sigma_z$ axis, and lies at angle of $-\phi$ from the $\sigma_x$ axis. This plane contains all of the states with satisfy the commutation condition. . . . .	118
5.6	<b>An expanded version of Figure 5.5.</b> The turquoise plane satisfies the commutation condition as before. We also plot two red planes (with dashed borders). Any states on these two planes will have a diagonal QFI matrix. Only the states on the intersections of a red plane with the turquoise plane satisfy both conditions, however, these states yield 0 information about one parameter. . . . .	119
5.7	<b>A plot of Equation (5.19)</b> for $\eta = \frac{1}{2}$ and any value of $\phi$ for the two-qubit state with $d_0 = \cos \theta_1$ , $d_1 = \sin \theta_1 \cos \theta_2$ and $d_{-1} = \sin \theta_1 \sin \theta_2$	126
5.8	<b>Incompatibility of optimal states.</b> Plots of the fisher information (normalised to a maximum of 1) of phase and dephasing for $N = 2$ . Recalling the form of the input states $ \psi_0\rangle = \sum_m d_m  J, m\rangle$ , we set $d_{-1} = d_1 = \frac{1}{\sqrt{2}} \cos(\theta)$ and $d_0 = \sin(\theta)$ , and vary $\theta$ along the $x$ -axis. Clearly the maxima as indicated by the dashed lines do not agree. . . . .	127
5.9	<b>Optimised plots for phase and dephasing.</b> In (a) we plot the quantum Fisher information for phase when all but one parameter $d_i$ is optimised and we vary this final parameter along the $x$ -axis. For the red line, these parameters are optimised for phase, with the blue line they are optimised for dephasing. Noting that the $y$ -axis begins at about a value of 2.45, it is apparent that there is not a huge difference between the optimal states for estimating phase and dephasing. (b) We now plot the values for dephasing in the same manner as for phase: for the blue line all but one parameter is optimised for dephasing and for the red, all but one is optimised for phase. We see similar results in that there is not a significant difference between optimal values. . . . .	128

## LIST OF FIGURES

---

5.10 An interferometer with loss in both arms and a phase shift $\phi$ in one of the arms. We study the case of symmetric loss where the transmissivities are given by $\eta_1 = \eta_2 = \eta$ . . . . .	129
---	-----



# Part I

# 1

## Introduction

At its inception, quantum theory was a tool for understanding phenomena which appeared often paradoxical when addressed with the existing tools of physics. With quantum mechanics, the mathematical paradoxes of what we now call classical mechanics (such as the ultraviolet catastrophe) were largely reduced to a perhaps more tenable class of intuitive paradoxes (such as the measurement problem). Aside from some scarce examples such as the photoelectric effect, the quantum world was far out of reach for manipulation and experimental observation. It was sufficiently remote that even great minds such as Schrödinger questioned the reality of what we now accept as canon [10],

*“We never experiment with just one electron or atom or (small) molecule. In thought experiments, we sometimes assume that we do; this invariably entails ridiculous consequences. In the first place it is fair to state that we are not experimenting with single particles anymore than we can raise ichthyosauria in the zoo.”*

Remarkably, these days we are raising our hypothetical ichthyosauria (though sadly, no literal ones); experiments are capable of generating single photons [11], manipulating individual ions [12], or addressing single-particle defects in crystals [13]. Simultaneously, we’ve experienced the rise of computation, and with it, the diminishment of scale: from computers filling rooms, to sitting in the palm of a hand. It is inevitable, in this sense, that with the miniaturisation of electronics we will hit a quantum wall.

It is thus necessary that we develop the capability to control quantum phenomena

---

in order to surpass the inevitable quantum obstacles we will encounter. However, in addition to the necessity for such control, it is even more desirable to saddle the ichthyosauria and harness quantum effects [14]. We can use quantum mechanics not merely as a tool for understanding reality, but also for exploiting it. Rather than just avoiding obstacles to the optimisation of existing technologies, we could develop entirely new ones.

This is the aim of quantum information. Quantum effects can in principle be used to enhance information processing [15], to increase the accuracy of measurements [16] and to enable unconditionally secure communication protocols [17]. Arguably, the first examples of quantum technologies are already widely in use, in the form of the transistor and the laser. On the flip side, these technologies might be seen not as exploiting quantum theory to an advantage for information processing in the ways mentioned. Lasers, for example, produce *coherent states*, which are considered to be archetypical ‘classical-like’ optical states [18].

In order to develop a deep understanding of the wide array of quantum technology proposals, it is important to grasp the underlying sources of enhancement over any classical counterparts. This problem has been tackled from numerous perspectives. On the very rigorous mathematical side, resource theories have been and are being developed, such as those of entanglement [19], non-locality [20], coherence [21] and even thermodynamics [22].

On the other hand, departing from this level of mathematical purity, it is a popular trend to study specific systems in attempt to identify what exactly the quantum resource they use is; from entanglement in universal quantum computation [23], to (allegedly) discord in DQC1 [24]. Often, also, the process is switched up and an operational setting is designed with the intention of defining a measure of some quantum correlation. One example of this is the interferometric power [25], which identifies discord-type correlations as the essential resource for a certain metrology scheme.

In this thesis, I do not consider the highly rigorous resource-theory approaches, but rather follow the path of trying to understand the functioning of particular physical systems. This work began with consideration of ghost imaging [26] – a system over which there had been some fair debate over the nature of its ‘quantumness’ – and treated it from an informational perspective. This led to considerable focus on intensity interferometry type schemes for some part of this thesis. The most well-known of these

## 1. INTRODUCTION

---

schemes is possibly the Hanbury Brown and Twiss stellar interferometer [27], which presented the first demonstration of second order coherence in the photon-counting regime.

Later, this work evolved to entirely different considerations of resources, in a study of multiple parameter quantum estimation. In this case, the resources are usually considered to be the number of qubits, for example, or the number of channel uses. Usually the goal of an estimation problem as considered in quantum metrology is to estimate some parameter with the optimal accuracy given limited resources (the problem would be rather trivial otherwise) [28, 29]. In multiple parameter estimation, this problem is generalised to optimal estimation of more than one parameter at the same time.

It is necessary to be very careful in our consideration of resources in quantum metrology. At times deceptively optimistic results have been attained due to poorly formulated problems [30, 31], and moreover, there have been some debates on probabilistic metrology schemes [32, 33], wherein we can discard some results by post-selection. It has been questioned if, given that one may discard resources that could be used in another way, these are of any real use [34]. Ultimately, a careful consideration of what constitutes a resource can make or break the utility of probabilistic schemes in general [6].

To precisely outline the structure of this thesis, chapter 2 introduces the essential tools of quantum information, starting with the basics of quantum mechanics, which it then combines with basic information theory to introduce the theory of quantum correlations. Chapter 3 is also an introductory chapter, this time for quantum optics and continuous variable quantum information. It has a significant focus on Gaussian states. Some of the more specific background may be contained outside of these two introductory chapters in the sections to which it is relevant.

Chapter 4 encompasses the work done on intensity interferometry schemes. It considers the Hanbury Brown and Twiss scheme, ghost imaging and quantum illumination. It examines classical and quantum correlations in all of these schemes. Chapter 5 lays out the work on multiple parameter metrology, with some significant background material as well as applications to a number of problems. Chapter 6 concludes the thesis.

## 2

# Toolbox

### 2.1 Quantum States and Observables

In quantum theory we consider two fundamental, intertwined objects: states and observables. There is significant work in the foundations of quantum mechanics to establish quite *how* these – states in particular – are fundamental [35, 36, 37], but we shall not concern ourselves with such matters. For our conceptual picture, it is sufficient to take the state as a complete description of the system and the observable as a property which we can measure with some macroscopic device.

In this section we shall only explicitly discuss finite-dimensional systems. Many of the ideas apply fairly readily to infinite-dimensional systems although there can be complications arising due to problems of boundedness, amongst other things. Although this thesis does deal a fair amount with infinite-dimensional systems, it primarily considers the Gaussian case, which is much simpler than the general one. That exposition will be relegated to the second chapter, as separating the analyses presents the path of least resistance. Most of the content in this chapter can be found in [38, 39], which are excellent resources for learning quantum information theory.

The most basic description of states is in terms of a ray  $|\psi\rangle$  in Hilbert space. This is known as a ket. Dual to this, we also define a bra  $\langle\psi|$ . Explicitly expressing the ket of an  $n$ -dimensional system in vectorial form  $|\psi\rangle = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ , we find the bra can be expressed as a row vector  $\langle\psi| = (c_1^*, \dots, c_n^*)$ , where the asterisk denotes the complex

## 2. TOOLBOX

---

conjugate. We can then define an inner product  $\langle\varphi|\psi\rangle$  on states with the usual matrix multiplication. States are normalised under this inner product,  $\langle\psi|\psi\rangle = 1$ , implying that  $\sum_i |c_i|^2 = 1$  but otherwise states are quite unrestricted.

We can decompose such a state into a *superposition* of states, by noting that a vector  $(c_1, \dots, c_n)^T$ , may be decomposed in terms of its basis vectors as  $\sum_{i=1}^n c_i |i\rangle$  where  $|i\rangle$  refers to the column vector with a 1 in the  $i^{th}$  position and 0 everywhere else. Clearly, as long as we are considering inner products, the choice of basis does not matter. If we take some change of basis enacted by a unitary  $U$ , such that  $|\psi'\rangle = U|\psi\rangle$  and  $|\varphi'\rangle = U|\varphi\rangle$  then  $\langle\varphi'|\psi'\rangle = \langle\varphi|\psi\rangle$ . The change of basis preserves the inner product.

Traditionally, the ‘overlap’  $|\langle\varphi|\psi\rangle|^2$  of two states is often explained as the probability of ‘jumping’ from one to the other, or as the transition probability. Perhaps more accurately, it characterises the indistinguishability of states via the *pure state fidelity* [40]. This implies that if we have simple measurement procedure which attempts to measure if a prepared state is  $|\psi\rangle$  or not  $|\psi\rangle$ , then the probability of getting a false positive if the actual state prepared was  $|\varphi\rangle$ , is given by pure state fidelity.

An observable,  $\hat{A}$  is generally considered to be a quantity represented by a self-adjoint operator, so  $\hat{A} = \hat{A}^\dagger$ . For finite-dimensional matrices the adjoint  $\hat{A}^\dagger$  is the conjugate-transpose of  $A$ . The expectation value of this observable can be calculated on a given state by the formula  $\mathbb{E}(\hat{A}) = \langle\psi|\hat{A}|\psi\rangle$ . We will often also use the notation  $\langle\hat{A}\rangle$  to denote the expectation value of  $\hat{A}$  where we do not expect it to cause any confusion with the bra-ket notation.

Of essential importance is the question of how a series of experiments measuring  $\hat{A}$ , on the same preparation of the state  $|\psi\rangle$  will break down into individual runs. This question relates to Born’s rule which is part of the fundamental postulates of quantum mechanics connecting individual measurement outcomes to observables. We will unravel the ideas behind this in a somewhat facile manner, which serves mainly to highlight the practical points.

Due to the self-adjointness of  $\hat{A}$ , we make use of the spectral decomposition. This states that for any normal operator on an  $n$ -dimensional Hilbert space  $\mathcal{H}$  (i.e. an operator for which  $\hat{A}^\dagger\hat{A} = \hat{A}\hat{A}^\dagger$ ), we can decompose it as  $\hat{A} = \sum_{i=1}^n \lambda_i |i\rangle\langle i|$ , where  $\lambda_i$  are the eigenvalues of  $\hat{A}$  and the  $|i\rangle$  form an orthonormal set of vectors. It is worth noting that the  $\lambda_i$  need not all be different, and that some may also be 0. We might

also notice that due to their orthonormality, we can write  $\sum_i |i\rangle\langle i| = \mathbb{1}$ , where  $\mathbb{1}$  is the identity on the Hilbert space  $\mathcal{H}$ .

This is clearly relevant to the calculation of  $\langle\psi|\hat{A}|\psi\rangle$ , which can be decomposed as  $\mathbb{E}(\hat{A}) = \sum_i \lambda_i \langle\psi|i\rangle\langle i|\psi\rangle$ . Given the fact that the  $\langle\psi|i\rangle\langle i|\psi\rangle$  sum to 1, we interpret them as probabilities. We shall write  $\text{pr}(i) = \langle\psi|i\rangle\langle i|\psi\rangle = |\langle i|\psi\rangle|^2$ . Then in an individual run of an experiment it can be found that outcome of the eigenvalue  $\lambda_i$  occurs with probability  $\text{pr}(i)$ . All of this is consistent with experimental evidence.

This forms the basis for speaking of measurement in general. We say that a measurement is described by a set of operators  $\{\hat{P}_i\}$  such that

$$\hat{P}_i\hat{P}_j = \hat{P}_i\delta_{ij}, \quad \hat{P}_i^\dagger = \hat{P}_i, \quad \sum_i \hat{P}_i = \mathbb{1}. \quad (2.1)$$

where  $\delta_{ij}$  is the Kronecker delta. This is not considered to be the most general type of measurement (although in another sense it is, a point we will clarify later). We can write that a measurement, described as above, on a state  $|\psi\rangle$  yields outcome  $i$  with probability  $\text{pr}(i) = \langle\psi|\hat{P}_i|\psi\rangle$ . Importantly, the post-measurement state will be given by  $|\psi_i\rangle = \frac{\hat{P}_i|\psi\rangle}{\sqrt{\text{pr}(i)}}$ . The probability in the denominator ensures proper normalisation.

The above formalism, though fundamental, is still quite basic and does not really go beyond what would be covered in elementary quantum mechanics. The first step of a more in-depth development requires a revisiting of the concept of a state.

### 2.1.1 Mixedness

It is clear from the above introduction that there is some inherently probabilistic aspect to quantum mechanics. Even in the simplest example of two-dimensional vector space, if we define  $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , where  $|0\rangle$  and  $|1\rangle$  are orthonormal vectors, and  $\hat{P}_1 = |0\rangle\langle 0|$ ,  $\hat{P}_2 = |1\rangle\langle 1|$ , we find the corresponding outcomes each occur with the probability of  $\frac{1}{2}$ .

Given the assumption of the state-vector as being a complete description of the system, the conclusion that quantum mechanics is inherently probabilistic is inevitable. However, this does not necessarily imply that every possible probabilistic aspect of the theory is contained in the above. Certainly, we might expect, as in classical mechanics, that we will find probabilities which arise due to incompleteness of our knowledge; that is, we do not know if a state is represented by one vector or another.

## 2. TOOLBOX

---

Thus, if we have a preparation procedure which produces  $|0\rangle$  and  $|1\rangle$  with equal probability, then we again expect to find that the measurement above yields a 50-50 chance of either outcome. That does not, however, imply that the states are the same: had we chosen to measure in the alternative basis  $\hat{P}'_1 = \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$  and  $\hat{P}'_2 = \frac{1}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)$ , the state  $|\psi\rangle$  would have produced a deterministic result, whereas the mixture would still give a  $\frac{1}{2}$  chance of each outcome. The states are physically distinguishable with the appropriate measurement procedure.

The probability of an outcome  $i$ , associated with a projector  $P_i$  where the preparation procedure probabilistically produces states  $|\psi_j\rangle$  with probabilities  $p_j$ , will be given by

$$\text{pr}(i) = \sum_j p_j \langle \psi_j | P_i | \psi_j \rangle \quad (2.2)$$

We are interested in writing this in a more explicit manner as a function of a *single* state and observable. By inserting simple resolutions of the identity: in particular sums of rank 1 orthonormal vectors on the whole space, we can write

$$p(i) = \sum_{j,k} p_j \langle \psi_j | \hat{P}_i | k \rangle \langle k | \psi_j \rangle \quad (2.3)$$

$$= \sum_{j,k} p_j \langle k | \psi_j \rangle \langle \psi_j | \hat{P}_i | k \rangle \quad (2.4)$$

$$= \text{Tr } \hat{\rho} \hat{P}_i \quad (2.5)$$

where the state is now described by the self-adjoint operator  $\hat{\rho} = \sum_j p_j |\psi_j\rangle \langle \psi_j|$  with trace 1, known as the *density matrix*. The trace operation of a matrix  $X$  is defined by  $\text{Tr } X = \sum_i \langle i | X | i \rangle$  where  $i$  is a complete set of orthonormal vectors, in effect, we take the sum of the diagonal entries of  $X$  (this is invariant under change of basis). Along with the self-adjointness and unit trace, we require the density matrix be positive, thus the defining criteria for a quantum state in the density matrix formalism are written as

$$\hat{\rho} = \hat{\rho}^\dagger, \quad \text{Tr } \hat{\rho} = 1, \quad \langle i | \hat{\rho} | i \rangle \geq 0 \forall |i\rangle \in \mathcal{H}. \quad (2.6)$$



Notably, the positivity condition actually implies self-adjointness, although these two conditions are conventionally stated separately. For an observable  $\hat{A}$  as described before, it is easy to see that the expectation value will be given by  $\text{Tr } \hat{\rho} \hat{A}$ . After a given measurement result corresponding to the operator  $\hat{P}_i$ , the post-measurement state will be given by  $\frac{\hat{P}_i \hat{\rho} \hat{P}_i}{\text{Tr } \hat{P}_i \hat{\rho} \hat{P}_i}$ .

The simple vector states we previously discussed can be written as  $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$  in this formalism, thus they are rank 1 density matrices. These are known as *pure states*; all other states are *mixed states*.

### 2.1.2 Evolution

Part of the intertwinement of state and observable is gleaned early in our learning of quantum theory when we discuss how systems change in time. The most fundamental quantum evolution is unitary evolution, which, as the name would indicate, is possible to describe by a unitary matrix.

Remarkably, there are two views of what it means for a system to evolve under a unitary: in the Schrödinger picture, it is the state which changes with respect to the evolution parameter (usually time). Alternatively it is also possible to view the state as fixed, and the operators as changing, in the so-called Heisenberg picture.

The equivalence of these two ideas arises immediately given the formalism we've already expounded. If we take a state vector in the Schrödinger picture, a unitary  $U$  acts on it by simple matrix multiplication,  $|\psi\rangle \rightarrow U|\psi\rangle$ . Thus, a density matrix evolves as  $U\hat{\rho}U^\dagger$  (this is true for mixed states too). Taking the cyclic property of the trace, we can see that if we wish to calculate the expectation of an observable on the evolved state, we get  $\langle\hat{A}\rangle = \text{Tr } U\hat{\rho}U^\dagger\hat{A} = \text{Tr } \hat{\rho}U^\dagger\hat{A}U$ . This actually reveals how observables evolve under the action of a unitary in the Heisenberg picture:  $\hat{A} \rightarrow U^\dagger\hat{A}U$ .

Now, it is important to note that this is not the only type of quantum evolution we consider. A unitary operation is always reversible, and a purely reversible theory would seem rather deficient given the presence of many irreversible processes in nature. The theory of *open quantum systems* [41] considers non-unitary evolutions as well, which are not reversible.

The form of general non-unitary evolutions can be axiomatically formalised [38]. Let  $\Lambda$  be a map from one set of density operators to another with the following properties

## 2. TOOLBOX

---

- Unit trace:  $\text{Tr } \Lambda(\hat{\rho}) = 1$ ,
- Convex-linearity:  $\Lambda(\sum_i p_i \hat{\rho}_i) = \sum_i p_i \Lambda(\hat{\rho}_i)$ ,
- Complete positivity:  $(\mathbb{1} \otimes \Lambda)(\hat{\rho}')$  is positive for any state  $\hat{\rho}'$ .

These are largely intuitive properties. The first requires the output to be a valid state. The second arises from the idea of a mixed state  $\hat{\rho} = \sum_i p_i \hat{\rho}_i$  as being produced by a process which yields  $\hat{\rho}_i$  with probability  $p_i$ . Thus the output will be  $\Lambda(\hat{\rho}_i)$  with probability  $p_i$ . The third also arises from the necessity of producing a legitimate density matrix as an output of the mapping; we require positivity of the density matrix, and thus the map itself is *positive*, meaning any positive input leads to a positive output. However, we also consider that the dynamics may only act on part of a total, larger state which is why we consider the composite mapping  $(\mathbb{1} \otimes \Lambda)(\hat{\rho}')$ , where the identity acts on the unaffected part of the state. When this too is positive for any  $\hat{\rho}'$ , we have *complete positivity*.

We call such maps completely positive trace preserving (CPTP) maps. It can be shown that the output of any CPTP map can be decomposed as  $\Lambda(\hat{\rho}) = \sum_i K_i \hat{\rho} K_i^\dagger$  where  $\sum_i K_i K_i^\dagger = \mathbb{1}$ , in what is known as the operator-sum representation. The operators  $K_i$  are called Kraus operators.

The complete positivity assumption links to an element we have not discussed – composite systems in quantum mechanics. If it's possible to have an operator acting on only part of a system, then we need to have the right language for discussing how to combine smaller systems into a larger one, and how to discuss the reduced components of the larger system. This is actually an extremely important topic, and understandably vital to the discussion of correlations in general. We thus talk about the framework for describing composite states next.

### 2.1.3 Composite Systems

As mentioned earlier, quantum mechanics occurs in Hilbert space. Thus when talking about a composite system, we need a composite space too. For this we need to use the tensor product; two Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  join to make a composite space  $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$ .

It is then quite easy to see how the discussion of non-interacting systems takes place in this setting. Simply enough, for uncorrelated systems (and pure states), we may write  $|\Psi\rangle_{12} = |\psi\rangle_1 \otimes |\varphi\rangle_2$ , where system 1 and 2 may be described by  $|\psi\rangle_1$  and  $|\varphi\rangle_2$  respectively. Likewise, operators acting independently on each system are combined by the tensor product  $\hat{A}_1 \otimes \hat{B}_2$ .

This yields the expected statistics for experiments on the system. Outcomes are statistically independent and the joint probability of outcomes  $i, j$  for the first and second systems respectively will be written as  $\text{pr}_{12}(i, j) = \text{pr}_1(i)\text{pr}_2(j)$ . Of course, we are also interested in correlated states; there are far more states in the space  $\mathcal{H}_{12}$  than just the uncorrelated ones.

We identify two classes of states. A *separable* state is any state which can be written as a convex combination of uncorrelated states,

$$\sum_i p_i \hat{\rho}_{i,1} \otimes \hat{\rho}_{i,2} \tag{2.7}$$

Measurements on each side of a separable state need no longer be statistically independent, since, if we again take the picture of a preparation procedure which produces pairs of states (each pair with a given probability), then the measurement statistics for each run would depend on which pair of states was produced, and since the state held by one party depends on the state held by the other, we get correlated measurement statistics.

There is a stronger form of correlation wherein the state cannot be factorised this way. Non-separable states are known as *entangled* states and are considered to be highly non-classical. In fact, entanglement is an essential component for universal quantum computation [23] and has been described, by none other than Schrödinger himself as [10], “*the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought*”.

An archetypical example of an entangled state is the *Bell state*,

$$\begin{aligned} |\phi^+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle) \\ &\equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \end{aligned}$$

## 2. TOOLBOX

---

where we henceforth adopt the concise convention for denoting the tensor product,  $|ijk\dots\rangle = |i\rangle \otimes |j\rangle \otimes |k\rangle \otimes \dots$ .

Entanglement complicates the notion of composite systems by undermining the language of speaking of systems as ‘separate’. For intuition, it is helpful to consider entanglement as a property occurring between spatially remote systems. One can then imagine remote observers performing measurements on these systems, so it certainly seems there is some kind of separation involved.

On the other hand, it is not possible to create entanglement by local operations and classical communication (LOCC), meaning that if two parties are remote, they cannot create entanglement. Thus it would appear the two ‘separate’ elements of the system which they are measuring have to have directly interacted at some point in the past at the very least.

The confusions are somewhat compounded by Bell’s inequalities which are violated by quantum theory. Bell-type inequalities are bounds on the degree of correlations we can measure given the assumption of local realism, although what this assumption means is not actually in itself necessarily clear [42, 43, 44]. One of the more famous Bell inequalities is the CHSH inequality, which can be formulated for binary systems as follows [45, 46]. Consider a system shared between two parties, Alice and Bob, each of whom can perform one of two measurements with outcomes of  $a, a'$  for Alice’s first and second measurement settings respectively and  $b, b'$  for Bob’s. Let the outcomes of any of these observables be  $\pm 1$ . Assume that the measurement outcomes are already predetermined by some ‘hidden variables’, then we can have on any experimental run, that  $a + a' = 0$ , in which case  $a - a' = \pm 2$ , or  $a + a' = \pm 2$  in which case,  $a - a' = 0$ . This allows us to define the correlation  $C \equiv (a + a')b + (a - a')b = \pm 2$ . From this we obtain

$$|\langle C \rangle| = |\langle ab \rangle + \langle ab' \rangle + \langle a'b \rangle + \langle a'b' \rangle| \leq 2 \quad (2.8)$$

which is violable by quantum mechanics, and the assumption that the properties  $a, a', b, b'$  have simultaneous definite values is thus incorrect.

Thus the language is somehow unclear. Is entanglement a correlation in a property which does not yet exist? If this is true and it only comes into being upon measurement of one system, this certainly seems to violate the idea of ‘separateness’ of each part of the system. However, the observers are certainly separate in our everyday understanding of

the word, and they also perform separate measurements too. Or perhaps the property does exist before measurement, but the system is highly non-local, as is explicitly the case in some hidden-variable theories like that of de-Broglie and Bohm [47, 48, 49].

Ultimately, we put philosophy aside and disregard this confusion, focussing mainly on the mathematical description at hand. As has been the case for most of this section, the description set forth is largely axiomatic, but extremely well-verified by experiment. The mathematical description of entanglement above is perfectly in alignment with experimental results.

### 2.1.4 Purification and Dilation

Despite the existence of mixed states and non-unitary evolution, the earlier description of pure states and unitary evolution as being *fundamental* was not a cavalier statement but actually a consequence of the fact that quantum theory can be axiomatised purely in terms of these [38, 50].

Just as we have rules for combining systems, it is also important to be able to consider parts of larger systems and this is the first step necessary for unravelling the above statement. Typically, we will picture a system and an environment; this may be quite literal, in that we are attempting to investigate the system in a laboratory, but are confounded by interactions with the larger world, but it can also be viewed simply as an abstract designation. The combined state of the system and environment will be called  $\hat{\rho}_{SE}$ , and the reduced states of system and environment independently,  $\hat{\rho}_S$  and  $\hat{\rho}_E$ .

It is again obvious that for quantum theory to be useful, we have to be able to describe the statistics of *what we can observe*, which may not be the whole system thus meaning the idea of reduced states is vital. Given a composite state of system and environment  $\hat{\rho}_{SE}$  (or any two-or-more party state), if we wish to only describe the system, then we have to trace over all degrees of freedom of the environment. More rigorously this *partial trace* is described by

$$\text{Tr}_E(|s_1\rangle\langle s_2| \otimes |e_1\rangle\langle e_2|) \equiv |s_1\rangle\langle s_2| \text{Tr}(|e_1\rangle\langle e_2|). \quad (2.9)$$

Letting this function be linear, we can then input density matrices into this function since any composite state will be a weighted sum of terms as above. Pursuant to

## 2. TOOLBOX

---

this, we write  $\hat{\rho}_S = \text{Tr}_E(\hat{\rho}_{SE})$ . Any operator acting solely upon the system will then yield the expected statistics using the usual formalism with  $\hat{\rho}_S$  as the state,  $\text{Tr} \hat{\rho}_S \hat{A} = \text{Tr} \hat{\rho}_{SE}(\hat{A} \otimes \mathbb{1})$ .

Using the Bell state as an example once again, and taking the standard quantum information practice of distributing the state between two parties named Alice ( $A$ ) and Bob ( $B$ ), we write

$$\begin{aligned}\hat{\rho}_A &= \text{Tr}_B\left(\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)(\langle 00| + \langle 11|)\right)_{AB} \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|).\end{aligned}$$

Thus in this case, Alice's reduced state  $\hat{\rho}_A = \frac{1}{2}\mathbb{1}$  will effectively produce random statistics. As a multiple of the identity matrix, it is *maximally mixed*. This is a defining feature of reduced density matrices of entangled states: roughly, the more mixed they are, the more entangled the initial state was. This connection will be explained further in the next section.

Building on the ideas of composite and reduced systems, we can introduce the notions of state purification and dilation. These are intuitively useful because they allow us to consider mixed and open quantum systems as parts of larger pure, unitarily evolving systems. Moreover, they are also of practical utility often allowing one to perform otherwise difficult calculations and optimisations [51].

The Bell state example revealed that we may find a mixed state as the reduced state of a pure one. Conversely, if we have a mixed state, it might be possible to find a pure state to which it acts as a reduced state (although this may really be a purely mathematical procedure, and not physically motivated). Indeed, *any* finite-dimensional mixed state allows a purification as a consequence of the *Schmidt decomposition*. This provides a very simple recipe: for  $\hat{\rho}_S = \sum_i \lambda_i |i\rangle\langle i|$ , let the environment be of the same dimension as  $\hat{\rho}_S$  such that we can easily write  $\psi_{SE} = \sum_i \lambda_i |i\rangle_S |i\rangle_E$ . The state has thus been purified.

Similar in spirit, we can show that all CPTP maps can be regarded as unitary operations on a larger system, using what is known as *Stinespring dilation*. In finite dimensions this is especially simple. Given an  $n$ -dimensional system, we require no more than  $n^2$  dimensions to describe the environment and moreover, the environment

starts in a pure state, uncorrelated with the system. Thus, for any CPTP map  $\Lambda$ , we can write

$$\Lambda(\hat{\rho}_S) = \text{Tr}_E(U^\dagger[\hat{\rho}_S \otimes |0\rangle\langle 0|_E]U), \quad (2.10)$$

where  $U$  is a unitary operating on both the system and environment.

### 2.1.5 Generalised Measurement

It was earlier mentioned that the description of measurement in terms of projectors (known as von Neumann measurement) is not the most general. This should be obvious, since a requirement such as idempotence necessitates that all measurements be repeatable, which is not in line with experimental evidence where we can have destructive measurements. However, much like with the case of mixed states and non-unitary evolution, any measurement can be rephrased as a projective measurement on a larger system. Clearly, we will lose some information if we then disregard part of that system. Thus these new measurements are suitable when we may not be interested in knowing the exact state of the subsystem after measurement.

Generalised measurements are known as positive operator valued measures (POVMs) and any POVM can be represented by a set of operators  $\{\hat{\Pi}_i\}$  satisfying

$$\hat{\Pi}_i^\dagger = \hat{\Pi}_i, \quad \sum \hat{\Pi}_i = \mathbb{1}, \quad \hat{\Pi}_i \geq 0. \quad (2.11)$$

Conversely any operators satisfying the above criteria are implementable POVMs. Unlike projective measurements, POVMs may be nonorthogonal and may not yield the same results upon repetitions of the measurement.

POVMs are connected with projective measurements via *Neumark's theorem*. Any  $n$ -dimensional POVM with elements  $\{\Pi_i\}$  can be written as a projective measurement on a larger space by appending an ancilla with an  $n$ -dimensional state space. Taking a pure state input  $|\psi\rangle$  to be measured, we initialise the ancilla (environment) in a state  $|0\rangle$  and then enact a unitary  $U$  defined by

$$U(|\psi\rangle_S \otimes |0\rangle_E) = \sum_{i=1}^n M_i |\psi\rangle_S |i\rangle_E. \quad (2.12)$$

## 2. TOOLBOX

---

Here  $M_i$  and  $M_i^\dagger$  are defined by  $\Pi_i = M_i^\dagger M_i$ . If we then define a projective measurement on the composite system  $P_i = \mathbb{1}_S \otimes |i\rangle\langle i|_E$ , we find that

$$\text{pr}(i) = \langle \psi |_S \langle 0 |_E U^\dagger P_i U | \psi \rangle_S | 0 \rangle_E = \langle \psi | \Pi_i | \psi \rangle, \quad (2.13)$$

thus achieving the POVM.

### 2.1.6 Qubits

To conclude the section on basic quantum mechanics, it is important to discuss the theory of qubit systems. Qubits are arguably the simplest quantum systems given that they live in the smallest non-trivial Hilbert space, and are the closest quantum analogy to a *bit*, as the fundamental binary units of classical information theory are called.

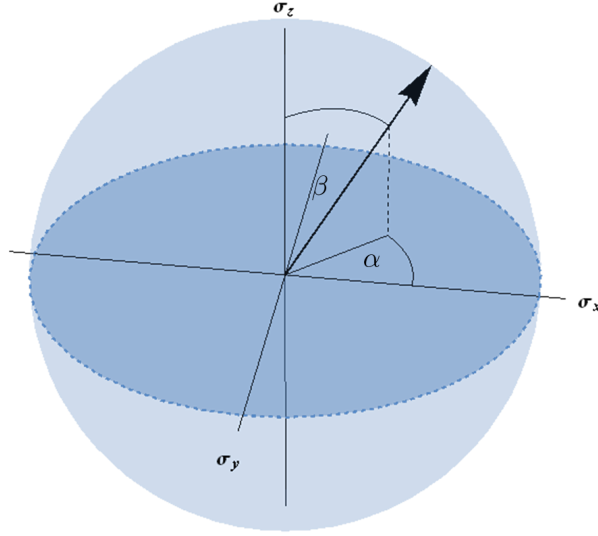
Qubits exist in a 2-dimensional Hilbert space, and we typically designate a *computational basis*, in which we denote the two orthogonal states by  $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . A general pure qubit state in this basis can then be written as  $|\psi(\alpha, \beta)\rangle = \begin{pmatrix} \cos \frac{\beta}{2} e^{-i\frac{\alpha}{2}} \\ \sin \frac{\beta}{2} e^{+i\frac{\alpha}{2}} \end{pmatrix}$ . It is thus evident that all pure states can be parametrised by two angles  $\alpha$  and  $\beta$ . A useful depiction of the space of qubit states is in terms of the Bloch sphere as in Figure 2.1.

It is almost unreasonable to speak of qubits without introducing the Pauli matrices. These generate the  $\text{SU}(2)$  algebra, which allows us to characterise all qubit unitaries with unit determinant. Additionally, they form a basis for all Hermitian  $2 \times 2$  matrices. Explicitly the Pauli matrices are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.14)$$

The Pauli matrices obey the commutation relation  $[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l$  and the anticommutation relation  $\{\sigma_j, \sigma_k\} = 2\delta_{jk}\mathbb{1}$ . The Levi-Civita symbol  $\epsilon_{jkl}$ , is 1 for all cyclic permutations of  $jkl$ , 0 if there are any repeated indices, and  $-1$  otherwise. Using the





**Figure 2.1: The Bloch sphere.** For a pure state, where the computational basis is taken to be the eigenbasis of the  $\sigma_z$  Pauli matrix, the two-angle parametrisation is represented as in the diagram. Alternatively, in the vector formalism, the origin is taken to be the centre of the Bloch sphere, and the  $x$ ,  $y$  and  $z$  components are given by the  $\vec{r}$  vector.

Pauli matrices, we may write the density matrix for a pure state as

$$\begin{aligned}\hat{\rho} &= |\psi(\alpha, \beta)\rangle\langle\psi(\alpha, \beta)| = \begin{pmatrix} \cos^2(\frac{\beta}{2}) & \cos(\frac{\beta}{2})\sin(\frac{\beta}{2})e^{-i\frac{\alpha}{2}} \\ \cos(\frac{\beta}{2})\sin(\frac{\beta}{2})e^{i\frac{\alpha}{2}} & \sin^2(\frac{\beta}{2}) \end{pmatrix} \\ &= \frac{1}{2} \left( \mathbb{1} + \begin{pmatrix} \cos(\beta) & \sin(\beta)e^{-i\frac{\alpha}{2}} \\ \sin(\beta)e^{i\frac{\alpha}{2}} & -\cos(\beta) \end{pmatrix} \right) \\ &= \frac{1}{2}(\mathbb{1} + \hat{n} \cdot \vec{\sigma}).\end{aligned}$$

Here  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$  and  $\hat{n} = (\sin \beta \cos \alpha, \sin \beta \sin \alpha, \cos \beta)^T$ . More generally, *any* density matrix may be written as  $\hat{\rho} = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma})$ , where  $\vec{r}$  is a vector such that  $|\vec{r}| \leq 1$ , as depicted in Figure 2.1.

As mentioned, the Pauli matrices also provide the basis for the generators of any single-qubit unitary with determinant equal to 1. Any such unitary can be written as,

$$e^{-i\frac{\theta}{2}\hat{n} \cdot \vec{\sigma}} = \cos \frac{\theta}{2} \mathbb{1} - i\hat{n} \cdot \vec{\sigma} \sin \frac{\theta}{2}. \quad (2.15)$$

It is also worth considering non-unitary evolutions on qubits. These are often characterised by bit-flips (where the  $|0\rangle$  and  $|1\rangle$  switch with a certain probability); phase

## 2. TOOLBOX

---

flips, where  $|0\rangle$  terms are unaffected but  $|1\rangle$  terms gain a  $\pi$  phase rotation (effectively a minus sign); and phase-bit flips, where both occur. There are a number of other types of *decoherence*, as these generally undesirable noise effects are known, which we won't discuss. To flesh out the example of the phase-flip, which is also known as dephasing or phase diffusion, we write the Kraus operators

$$K_1 = \sqrt{\frac{1+\eta}{2}} \mathbb{1}, \quad K_2 = \sqrt{\frac{1-\eta}{2}} \sigma_z. \quad (2.16)$$

The resultant density matrix, starting from an input  $\hat{\rho} = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix}$ , will be

$$\hat{\rho}' = \begin{pmatrix} a & \eta c \\ \eta c^* & b \end{pmatrix}. \quad (2.17)$$

Thus, the off-diagonals shrink, which corresponds to a contraction of the Bloch sphere around the equator (resulting in a prolate spheroid), where  $\eta$  is the radius of the contracted sphere.

## 2.2 Correlations

### 2.2.1 Entropy

The idea of entropy is indispensable for the discussion of correlations. Intuitively, this should be of no surprise at all: if we associate entropy with disorder and correlations with order, then one should arise from the other. It is useful to begin with an idea of entropy not from physics, but from information theory. Expounded in Shannon's seminal work [52], and identified as entropy by von Neumann, the *Shannon entropy* of a discrete probability distribution  $p_a$  associated with the outcomes  $a$  of some random variable  $A$  is defined as

$$H(A) = - \sum p_a \ln p_a. \quad (2.18)$$

This can be viewed as the expectation value of  $-\ln p_a$ , which following Shannon can be dubbed the *information content* of outcome  $a$ . The smaller the probability of an outcome, the more information it carries, we in some sense learn more when it occurs – we are more surprised. We thus already see an explicit link between entropy and information, defined in this sense.

We can question how to extend this quantity from classical probability distributions, to quantum systems. It is standard to identify the quantum equivalent of a probability distribution with the density matrix  $\hat{\rho} = \sum_a p_a |a\rangle\langle a|$ . Fixing the kets  $|a\rangle$  to be orthonormal, we can precisely define an equivalent to the Shannon entropy

$$\begin{aligned} S(\hat{\rho}) &= -\sum p_a \ln p_a \\ &= -\text{Tr } \hat{\rho} \ln \hat{\rho}. \end{aligned} \tag{2.19}$$

This is known as the von Neumann entropy and when the density matrix is decomposed in orthonormal form, can be seen to be the information content of the classical distribution of states in the ensemble. This is quite remarkable because it indicates that when  $\hat{\rho}$  is a pure state, i.e.  $p_a = 1$  for one value of  $a$  and 0 for all others, the entropy is 0. This may seem somewhat surprising, since measurement of pure states is still often associated with some inherent uncertainty. However, if given the correct basis to measure in, then in principle there is no uncertainty at all. Our uncertainty on pure state measurement might be said to arise from a lack of knowledge of the correct basis to measure in. However, we have defined the von Neumann entropy as the entropy of the measurement probability distribution *when measuring in the same basis* as the orthonormal decomposition of the density matrix, thus circumventing any such uncertainty.

On this note, it is useful to recognise another definition of disorder in quantum states. Ordinarily simpler than the von Neumann entropy to calculate, the *purity* of a state is defined as

$$\begin{aligned} \mu(\hat{\rho}) &= \text{Tr } (\hat{\rho}^2) \\ &= \sum_a p_a^2, \end{aligned} \tag{2.20}$$

where we obtain the second line by decomposing the density matrix in terms of orthonormal eigenvectors once again. This ranges from  $\frac{1}{n}$  at minimum, to 1, at its maximum, where  $n$  is the dimension of the Hilbert space. Conversely, we can define the linear entropy by  $S_L(\hat{\rho}) = \frac{n}{n-1}[1 - \mu(\hat{\rho})]$ . This ranges from 0 to 1 and is actually a first order approximation to the von Neumann entropy.

## 2. TOOLBOX

---

### 2.2.2 Entanglement

When introducing entanglement, we considered the example of a Bell state where the partial trace over either subsystem yields a maximally mixed state in the other. This property is not coincidental and is actually a consequence of the position of Bell states as *maximally entangled* states of two qubits.

For pure states, any correlation at all indicates entanglement. The fact that the information is shared across both parties prevents either party from having a well defined state on its own, and thus any local observable will reveal mixedness.

This actually provides one of the most consistent definitions for quantifying the entanglement  $\mathcal{E}$  of a pure bipartite system  $\hat{\rho}_{AB}$ . Whereas there are numerous entanglement measures, a great many reduce to the same simple expression on pure states

$$\mathcal{E}(\hat{\rho}_{AB}) = S(\hat{\rho}_A) = S(\hat{\rho}_B). \quad (2.21)$$

It is very important to note that this definition only applies to pure state entanglement. To quantify entanglement for mixed states we often need to consider the so-called convex roof extension [53]. One such example is the *entanglement of formation*, which for a mixed state  $\hat{\rho}$  considers the minimum average entanglement of an ensemble of pure states constituting  $\hat{\rho}$ .

### 2.2.3 Mutual Information

One of the broadest ideas of correlation, is that of the *mutual information*, which we will also refer to as the total correlations. Recalling the picture of Shannon entropy as average information content, the mutual information quantifies the shared information content between two systems. Yet again, the definition of this quantity is very intuitive.

$$\mathcal{I}_C(A : B) = H(A) + H(B) - H(A, B), \quad (2.22)$$

where  $H(A)$  and  $H(B)$  are the entropies of the marginal probability distributions  $p_a$  and  $p_b$  and  $H(A, B)$  of the joint distribution  $p_{a,b}$ . This is to be understood as a positive quantity, which is only 0 when there is no information shared by the two parties involved.

In the picture of Shannon entropy as average information content, we can see that this expresses that the expected information of the composite system is less than that

of each subsystem taken independently. This can be understood as a consequence of counting information redundantly: there is shared information which we count twice if we consider the systems independently i.e. *mutual information*. On the other hand, in the picture of entropy as disorder, it can be understood that the disorder of the composite system is less than that of the sum of the subsystems separately, this *lack* of disorder can immediately be identified with correlations. Both of these intuitive pictures are quite useful and help to understand the connection between entropy and information.

We can identify the quantum mutual information of a state by an almost exact analogy, once again substituting the Shannon entropy for the von Neumann entropy,

$$\mathcal{T}_Q(A : B) = S(\hat{\rho}_A) + S(\hat{\rho}_B) - S(\hat{\rho}_{AB}). \quad (2.23)$$

There is, however, a small practical difference. While, one can in principle measure the classical mutual information by purely separate measurements on each party, this is not the case for the quantum mutual information. Recalling that the von Neumann entropy corresponds to the entropy of the probability distribution arising from a measurement in the orthonormal basis of the density matrix, it may be the case that the orthonormal decompositions of the subsystems  $\hat{\rho}_X$  and  $\hat{\rho}_Y$  are not compatible with the orthonormal decomposition of  $\hat{\rho}_{XY}$ , which may not even be factorisable if the state is entangled.

This difference is not normally regarded as significant in of itself, however, active consideration of measurement is actually crucial when it comes to attempting to differentiate the quantum and classical components of mutual information.

### 2.2.4 Classical and quantum correlations

Given that the mutual information may not always be assessable from the outcomes of repeating one pair of separable measurements, we are interested in asking when it is the case that it *is* possible to do so. This would allow us to identify a class of states which behave in an effectively classical way. We desire that the measurements be factorisable (and thus we can expect separability to be necessary) but on top of that we also need the state to factorise as

$$\hat{\rho}_{AB}^{(CC)} = \sum_{ij} p_{ij} |i\rangle\langle i|_A \otimes |j\rangle\langle j|_B, \quad (2.24)$$

## 2. TOOLBOX

---

where  $|i\rangle_A$  and  $|j\rangle_B$  are orthonormal states on Alice and Bob's sides respectively. Under this condition, the same states will also arise in the orthonormal decomposition of the reduced density matrices and the quantum mutual information will be deducible purely by Alice and Bob measuring the marginals independently (in the appropriate basis) and then collating their data. These states are known as classical-classical (CC) states.

What will interest us more than these CC states, are quantum-classical (QC) states, which are in a manner of speaking classical with regards only to measurements on one side. It is with these sorts of states that the first quantifiers of 'quantum correlations' in separable states began to be explored [54, 55], where previously quantum correlations had been considered to be synonymous with entanglement [56]. Such QC states can be written as

$$\hat{\rho}_{AB}^{(QC)} = \sum_i p_i \hat{\rho}_{i,A} \otimes |i\rangle\langle i|_B. \quad (2.25)$$

The way the one-sided classicality of these states is usually described, is that there exists a measurement on Bob's system which does not disturb the overall state, but none on Alice's system with the same property. Of course, it is also possible to define a state with the reversed property (a CQ state).

It is desirable to formalise a quantifier for the level of this disturbance in a general quantum state, separable or otherwise. To do this, we define a new expression for mutual information which, when dealing with probability distributions in the classical domain, is completely analogous to  $\mathcal{T}_C(A : B)$ , but for which the quantum variant may differ from  $\mathcal{T}_Q(A : B)$ . Using Bayes' rule for conditional probabilities, it is possible to rewrite Equation (2.22) as

$$\mathcal{J}_C(A|B) = H(A) - H(A|B), \quad (2.26)$$

where  $H(A|B)$  is the conditional entropy of Alice dependent on Bob,  $H(A|B) = -\sum_{a,b} p(a,b) \log p(a|b)$ , where  $p(a|b)$  is the conditional probability of Alice observing  $a$  conditional on Bob's outcome  $b$ . In the classical case the conditional entropy is uncontroversial, mathematically, due to Bayes' rule, and also from the perspective of physical intuition where we can calculate all probabilities, whether joint or conditional, by using separate measurements on both systems.

In the quantum case, however, things are not so simple, and we need to consider the role of measurement very explicitly in the picture; to deduce the conditional state of Alice, we need to have measured Bob. Given a projective measurement  $\{\mathbb{1} \otimes \Pi_i\}$ , and an outcome corresponding to the element  $\mathbb{1} \otimes \Pi_i$ , the conditional state of Alice is given by  $\hat{\rho}_{A|i} = \frac{\text{Tr}_B \hat{\rho}_{AB}(\mathbb{1} \otimes \Pi_i)}{p_i}$ , where  $p_i = \text{Tr} \hat{\rho}_{AB}(\mathbb{1} \otimes \Pi_i)$ . The conditional measurement entropy is then given by  $S_{\{\Pi_i\}} = \sum_i p_i S(\hat{\rho}_{A|i})$ . Just as we technically consider the projective measurement which minimises the entropy of the output probability distribution with the standard von Neumann entropy, we will do likewise with the conditional one and thus [55].

$$\mathcal{J}_Q(A|B) = S(\hat{\rho}_A) - \inf_{\{\Pi_i\}} S_{\{\Pi_i\}}(\hat{\rho}_{A|B}). \quad (2.27)$$

This quantity is equivalent to  $I_Q(A : B)$  when the state is QC. This leads us to define the *quantum discord* by [54]

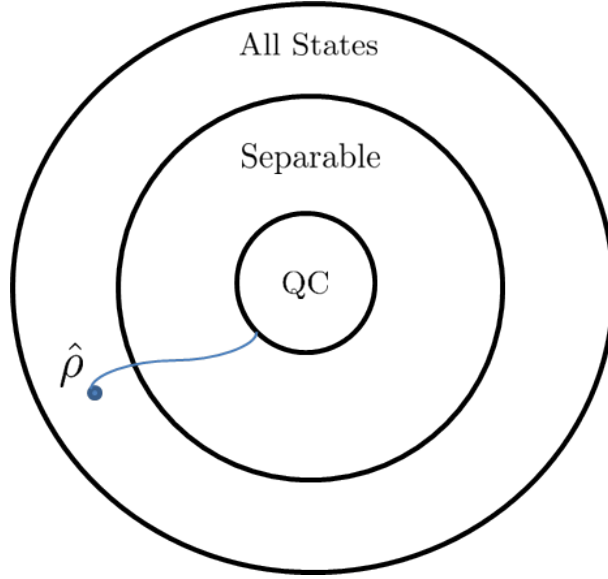
$$\mathcal{D}(A|B) = \mathcal{T}_Q(A : B) - \mathcal{J}_Q(A|B). \quad (2.28)$$

The quantum discord is often regarded as a type of quantum correlation. Strictly speaking, however, there are some flaws in this designation, given that the discord is asymmetric, generally  $\mathcal{D}(A|B) \neq \mathcal{D}(B|A)$ , which seems to defy conventional standards for what one considers to be a correlation. Moreover, it is possible to create discord with local (non-unitary) operations [57]. Nevertheless, it certainly captures some element of quantumness in bipartite systems; it is the nonorthogonality present in one half of the decomposition of the shared state which causes a non-zero discord. Such nonorthogonality is clearly not possible in the classical theory of probability. We will consider the quantity  $\mathcal{J}_Q(A|B)$  to be representative of the classicality of a system [55].

Aside from the entropic definition of discord, there are also a number of other definitions [58]. For its intuitive power, we shall mention the geometric picture of quantum correlations, although we do not use it in this thesis. In formulating the geometric picture, we classify states in terms of their quantumness as in Figure 2.2. Then with a suitable metric, we can quantify the distance between, for example an entangled state and the nearest separable state, or the nearest QC state. The magnitudes of these distances serve as quantifiers of correlations – of entanglement in the former case

## 2. TOOLBOX

---



**Figure 2.2: Geometric Quantum Correlations.** An intuitive notion of quantum correlations is gained from classifying states as in the diagram. For some state  $\hat{\rho}$ , we measure the distance to the nearest point in the desired set (QC for discord, for example). The geometric picture above is not representative of the actual shapes or size of the sets in any reasonable sense, but purely for demonstrative purposes [8].

and discord-type correlations in the latter. This provides a neat, unified picture of correlations and correlation-like quantities.

### 2.2.5 Other entropies

To conclude this section, I will discuss a different type of entropy to the Shannon and von Neumann ones. The Rényi-2 entropy will prove to be especially useful in the Gaussian setting for its computational simplicity and it will be used to analyse some of the later work of the thesis.

Starting again with the Rényi entropy for classical probability distributions, we define the Rényi- $\alpha$  entropy as [59],

$$H_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_{x=1}^n p_x^\alpha\right). \quad (2.29)$$

The index  $\alpha$  indicates that the Rényi entropy is actually a family of entropies, where  $\alpha \in [0, \infty)$ , not including  $\alpha = 1$ . As  $\alpha \rightarrow 1$ , we recover the Shannon entropy. The



quantum Rényi entropy is defined almost identically.

$$S_\alpha(\hat{\rho}) = \frac{1}{1-\alpha} \log \text{Tr } \hat{\rho}^\alpha. \quad (2.30)$$

This tends to the von Neumann entropy in much the same way the classical Rényi entropy tends to the Shannon entropy. We can also see that for the special case of  $\alpha = 2$ , we get the negative log of the purity  $S_2(\hat{\rho}) = -\log(\text{Tr } \hat{\rho}^2)$ .

Given this simplicity, it might seem curious why the von Neumann entropy holds a very privileged position in quantum information while the Rényi entropy does not. The understanding of this point arises from the *strong subadditivity inequality*, which states that for a *tripartite system*

$$S(\hat{\rho}_{ABC}) + S(\hat{\rho}_B) \leq S(\hat{\rho}_{AB}) + S(\hat{\rho}_{BC}). \quad (2.31)$$

The importance of this is not at all obvious at first, but it is extremely powerful when it holds, which it does for the von Neumann entropy but not generally for other entropies [38, 39]. A trove of desirable and informative properties of entropic quantities spring up as a result of the inequality, such as the fact that discarding systems doesn't increase mutual information, and the fact that conditioning reduces entropy and even that the conditional quantum mutual information is positive. It is very important. We shall see in the following introduction to continuous variables that the Rényi-2 entropy *does* satisfy this inequality when calculated on Gaussian states [60].

### 3

## Continuous Variables

In classical information theory, bits are fundamental. Similarly, in quantum information, qubits often take centre stage. However, qubits have some distinct disadvantages over their classical counterparts, most especially their fragility: in order to see quantum effects, we tend to encode qubits onto extremely tiny, extremely cold systems. A large part of quantum information aims to establish the groundwork for robust, easily manipulable qubits.

However, there is an alternative non-qubit approach to quantum information. Rather than use these discrete qubits, we encode information onto what we might think of as ‘analogue states’. A recent work has shown that any qubit protocol can be translated into a continuous variable one [61]. Moreover, it has been known for some time that universal quantum computation can be performed on continuous variable systems [62]. Even *more* significantly, this can be done in large part, but not exclusively, with Gaussian states and Gaussian measurements (unfortunately, Gaussianity must be compromised either in the states or in parts of the measurement phase but we can still go a long way with Gaussianity alone [62]).

In the following section I will introduce the general phase-space framework necessary for understanding continuous-variable quantum information, which I have used heavily throughout this thesis. Subsequently, I’ll focus on Gaussian states in particular, which have a beautiful, simple framework for use in the phase-space picture [4, 63, 64, 65, 66, 67].

### 3.1 Quantum Optics

Although continuous variable quantum information can be implemented in a variety of settings, I shall focus on optical schemes. These are not only very well investigated and characterised, but also convenient for experimental implementation. The practical tools of quantum optics are very well developed and simple to use. As a first step to quantum optics, we consider the quantisation of the electric field. Necessary for this, are the bosonic mode operators, also commonly known as creation and annihilation operators,  $\hat{a}$  and  $\hat{a}^\dagger$ . These operators obey the commutation relation,

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{k,k'}$$

where we implicitly assume  $\hbar = 1$  and  $\delta_{k,k'}$  is the Kronecker delta. These operators are necessary for the quantisation of fields (e.g. electric and magnetic) which are integral to the formalism of classical optics. Glossing over the detail, they arise from an analogy with the Fourier decomposition of the classical field, where a vector potential  $\mathbf{A}(\mathbf{r}, t)$  is decomposed into Fourier components, such that we might say  $\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}} \sum_{\lambda=1}^2 \mathbf{e}_{\mathbf{k}\lambda} A_{\mathbf{k}\lambda}(\mathbf{r}, t)$  where  $\lambda$  is a polarisation index, and  $\mathbf{k}$  represents the wave vector. After decomposing the field into conjugate components,  $\mathbf{A}_{\mathbf{k}\lambda}(\mathbf{r}, t) = \mathbf{A}_{\mathbf{k}\lambda} \exp(-i\omega_k t + i\mathbf{k} \cdot \mathbf{r}) + \mathbf{A}_{\mathbf{k}\lambda}^* \exp(i\omega_k t - i\mathbf{k} \cdot \mathbf{r})$  and subsequently replacing the components  $\mathbf{A}_{\mathbf{k}\lambda}$  with  $\hat{a}_{\mathbf{k}\lambda}$  and  $\mathbf{A}_{\mathbf{k}\lambda}^*$  with  $\hat{a}_{\mathbf{k}\lambda}^\dagger$ , this allows us to express the quantised Fourier decomposition of the field as [68, 69]

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}} \sum_{\lambda} \hat{a}_{\mathbf{k}\lambda} e^{-i\omega_k t + i\mathbf{k} \cdot \mathbf{r}} + \hat{a}_{\mathbf{k}\lambda}^\dagger e^{i\omega_k t - i\mathbf{k} \cdot \mathbf{r}}. \quad (3.1)$$

We can rewrite the mode operators in terms of *quadrature operators*,  $\hat{a}_k = \frac{1}{\sqrt{2}}(\hat{q}_k + i\hat{p}_k)$  and  $\hat{a}_k^\dagger = \frac{1}{\sqrt{2}}(\hat{q}_k - i\hat{p}_k)$ , with  $[\hat{q}_k, \hat{p}_{k'}] = i\delta_{k,k'}$ . We can define quadrature eigenvalues and eigenfunctions so that  $\hat{q}_k|q\rangle_k = q_k|q\rangle_k$ , though the explicit form of these is not particularly important. The quadrature operators, as canonically conjugate variables, can be considered as analogies to position and momentum operators.

A popular basis for describing optical states is the photon-number basis, also known as the Fock basis, where each element  $|n\rangle$  describes states of precise photon number,  $n$ . Upon these states, the creation and annihilation operators act as

### 3. CONTINUOUS VARIABLES

---

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \quad (3.2)$$

This allows us to define the number operator  $\hat{a}^\dagger\hat{a}$ , for which the Fock states  $|n\rangle$  are eigenstates such that  $\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle$ .

Equally important to this (more so, for the purposes of this thesis), are the coherent states, which in a sense also form a basis for describing optical states [70]. Coherent states are defined as the eigenstates of the mode operators  $\hat{a}$ . This makes it immediately obvious that we require an infinite-dimensional Fock space representation to describe the coherent states: if there were a maximum term  $|n\rangle$ , then clearly this component would be removed by the action of the annihilation operator and the state would thus change. It turns out that states which satisfy this are given by the pure states which produce a Poissonian counting distribution, meaning that  $\text{pr}(n) = \text{Tr } \hat{\rho}\Pi_n$  are Poissonian in  $n$  where  $\Pi_n = |n\rangle\langle n|$ . Pure states with this property are written as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (3.3)$$

These states form an overcomplete basis, satisfying the completeness condition  $\frac{1}{\pi} \int |\alpha\rangle\langle\alpha| d^2\alpha = 1$ , though no two states are orthogonal  $|\langle\alpha|\beta\rangle|^2 = e^{-|\alpha-\beta|^2}$ . Coherent states are as close to a deterministic classical state as we can really come in quantum optics; that is, they are *almost* a point in phase space, as will be described in more detail in the next section. Coherent states are certainly not the only classical-like states (although they have the unique distinction of being the only *pure* classical-like states). We can also identify classical-like mixed states, the most important of which are the *thermal states*. Together, all the classical-like states can be described in what is known as semi-classical optics, where the light field is described by classical optics but the detectors are quantum mechanical. This moves the fundamental source of noise from the e.g. Poissonian statistics of the light field, to the shot noise in the detectors themselves but produces mathematically identical results. In fact, in the quantum optical context we've been discussing, admitting a semi-classical description is precisely what we mean when speaking of classical-like states, and the states which allow this are well-characterised by the behaviour of their Glauber-Sudarshan  $P$ -distribution which

will also be expanded upon in the coming section. In the context of classicality in optical states, we will use the terms semi-classical, classical and classical-like interchangeably.

Coherent states are minimum uncertainty states in the quadratures, meaning  $\text{var}(q) \text{var}(p) = \frac{1}{4}$ , where  $\text{var}(x)$  refers to the variance of  $x$ , but other classical-like states will not be, due to their mixedness. We especially focus on thermal states, which are isotropic in phase space and have no coherences in the Fock basis. They can be decomposed according to the distribution

$$\hat{\rho}_{th} = \frac{1}{\bar{n} + 1} \sum_n \left( \frac{\bar{n}}{\bar{n} + 1} \right)^n |n\rangle\langle n|, \quad (3.4)$$

where  $\bar{n}$  is the mean photon number. We accordingly find that the uncertainty relation for these states reads  $\text{var}(q) \text{var}(p) = (\frac{1}{2} + \bar{n})^2$ . Thermal states occur quite naturally, and may be produced by black body sources.

To conclude this section, we will discuss one example of a state which is not classical-like. Thus far, the states considered have been isotropic, with  $\text{var}(q) = \text{var}(p)$ , however, there exist states with asymmetric quadrature variances. Most notably, we can find minimum uncertainty states with this quality, where one quadrature variance is ‘squeezed’ and the other is ‘anti-squeezed’, to compensate (since otherwise we’d violate the Heisenberg uncertainty principle). The Fock state decompositions of such states are interesting as they have no odd photon-number terms. We can write a squeezed vacuum state as,

$$|\psi\rangle_{sq} = \frac{1}{\sqrt{\cosh r}} \sum_n (-\tanh r)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle, \quad (3.5)$$

for some squeezing parameter  $r$ .

All the states considered above are known as Gaussian states. They form vital parts of the machinery for continuous variable quantum information and it is thus useful to explore the powerful techniques which exist for performing calculations with such states. On the way to doing so, we must first tackle the concepts of quasiprobability.

## 3.2 Quasiprobability

Probability theory was first formalised by Kolmogorov [71] and can be condensed to three axioms. For a measure space  $(\Omega, \mathcal{S}, \text{pr})$  [72],

### 3. CONTINUOUS VARIABLES

---

1.  $\text{pr}(E) \geq 0$  and  $\text{pr}(E) \in \mathbb{R}$  for all  $E \in \mathcal{S}$ . This says the probability of an event,  $E$ , is a real number greater than or equal to 0.
2.  $\text{pr}(\Omega) = 1$ . This says the probability that any event occurs is 1, i.e. some event must occur.
3.  $\text{pr}(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \text{pr}(E_i)$ , for  $E_i \cap E_j = \emptyset \quad \forall i, j$ . This ensures that the probabilities of mutually exclusive events are additive.

This is far more rigorous and abstract than we require, and this thesis certainly does not explicitly consider the measure-theoretic formulation of probability theory. However, it is useful to keep in mind an informal idea of the requirements. This will allow us to see which are to be sacrificed when considering *quasiprobability* functions. Ultimately, we desire positivity of probabilities, reality of probabilities, summation of mutually exclusive events and the equivalence to unity of the whole event space.

We will continue with some ideas from the classical theory of probability as they will occur also in quasiprobability theory. For starters, we take the formulae for the normalisation of probabilities and for calculating marginal probabilities to find,

$$\int_{x \in X} \text{pr}(x) dx = 1, \quad \int_{y \in Y} \text{pr}(x, y) dy = \text{pr}(x).$$

Here  $x$  and  $y$  refer to individual events, and  $X$  and  $Y$  to the entirety of the event space.

Another absolutely essential ingredient, especially for the understanding of Gaussian states, is that of *moments*. The moments are quite simply defined by  $\nu_r \equiv \langle x^r \rangle = \int_x x^r \text{pr}(x) dx$  where  $r$  is some integer; for example when  $r = 1$  we have the first moment, which is the mean. More essential than the moments though, are the central moments,

$$\mu_r \equiv \langle (x - \langle x \rangle)^r \rangle = \int_x (x - \langle x \rangle)^r \text{pr}(x) dx.$$

We will often use the notation  $\Delta x = x - \langle x \rangle$ . The most important of these is the second central moment, which is simply the variance. It is often denoted by  $\sigma^2$  or  $\Delta^2 x$ . A common reparameterisation is to take  $y = \frac{x - \langle x \rangle}{\sigma}$ ,  $y$  is then said to be in standard form.

To deal with moments in a powerful way, we use generating functions. In particular, the moment generating function is defined by  $M(\xi) \equiv \langle e^{\xi x} \rangle = \int e^{\xi x} \text{pr}(x) dx$ . From this, it is always possible to find any of the moments by noting that  $\langle e^{\xi x} \rangle = \langle 1 + x\xi + \frac{x^2}{2!}\xi^2 + \dots \rangle$

and subsequently we see that  $[d^r \frac{M(\xi)}{d\xi^r}]_{\xi=0} = \langle x^r \rangle$ . The central moments can similarly be calculated by replacing  $x$  with  $x - \langle x \rangle$ .

More important than this moment generating function, is the complex moment generating function  $C(\xi) \equiv M(i\xi)$ , which is significant enough to deserve its own line.

$$C(\xi) = \int e^{i\xi x} \text{pr}(x) dx. \quad (3.6)$$

This is known as the characteristic function and, unlike the moment generating function, it always exists since it effectively enacts a Fourier transform of the probability distribution, and the Fourier transform of a square integrable function is always a valid function.

We are now sufficiently equipped to enter the theory of quasiprobability. Though there are actually innumerable quasiprobability functions, we shall consider only three, which are quite closely related to each other, and are the most well-known. The first quasiprobability function to be discovered was the *Wigner function*. This was derived in 1932 by Eugene Wigner [73] although it's full use only gained appreciation in the 1940s with the work of J. E. Moyal [74], but not without significant resistance as evidenced by a quotation of Dirac,

*"I think it would be a good idea to have your work discussed, if you don't mind possible heavy criticism" [75].*

The 'quasi-' part of the Wigner function is primarily expressed in that it may be negative. It is possible to remove this issue by convolving the function with a Gaussian, as will be more explicitly stated later. This produces a well-behaved function, the Husimi  $Q$ -distribution [76], which was developed in 1940 in Japan, but unknown to Western scientists at the time. Nevertheless, this distribution violates the third axiom of probability and thus is indeed a quasiprobability distribution.

Finally, in the opposite direction from the  $Q$ -distribution is an 'unsmoothed' distribution, the Glauber-Sudarshan  $P$ -distribution [18, 77]. This function is a bit of a wild-card in that it is liable to be extremely poorly behaved.

### 3. CONTINUOUS VARIABLES

---

#### 3.2.1 Phase Space

The proper context for framing the quasiprobability distributions is as distributions on quantum phase space. A classical particle is fully determined by a point in phase space (assuming we know all relevant forces acting upon it). However, there still may be an epistemological uncertainty, in that we lack knowledge of the initial configuration and can only describe it by some probability distribution  $\text{pr}(q, p)$ . In this case, we often work with the moments of any function which depends on the canonically conjugate variables  $q$  and  $p$ , as these are especially helpful for describing the properties of the system.

In the quantum case, it is meaningless to consider a point in phase space due to the uncertainty relation: no state with well defined position and momentum can exist. This implies that we indeed need some sort of distribution on the phase-space to describe a quantum state.

We begin with the Wigner distribution, for which we require a single postulate [78, 79], which is that the quasiprobability distribution is a real distribution which produces the correct marginals. Thus, we expect each quadrature distribution to be contained in the quasiprobability distribution  $W(q, p)$ , via  $\text{pr}(q) = \int_{-\infty}^{\infty} W(q, p) dp$  and  $\text{pr}(p) = \int_{-\infty}^{\infty} W(q, p) dq$ . The postulate we use is even slightly more general than this.

$$\begin{aligned}\text{pr}(q, \theta) &= \langle q | U(\theta) \hat{\rho} U^\dagger(\theta) | q \rangle, \\ &= \int_{-\infty}^{\infty} W(q', p') dp',\end{aligned}$$

where  $q' = q \cos \theta - p \sin \theta$  and  $p' = q \sin \theta + p \cos \theta$  are rotated quadratures and  $U(\theta) = e^{i\theta \hat{a}^\dagger \hat{a}}$  is a phase rotating operator,  $\hat{\rho}$  is the density matrix of the state and we have only considered the single mode case.

To derive the explicit form of  $W(q, p)$ , we consider the Fourier transformed Wigner function and the characteristic function of the marginals

$$\tilde{W}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(q, p) e^{-iuq - ivp} dq dp, \quad (3.7)$$

$$\tilde{\text{pr}}(\xi, \theta) = \int_{-\infty}^{\infty} \text{pr}(q, \theta) e^{-i\xi q} dq. \quad (3.8)$$



Invoking the postulate, one can show that  $\tilde{p}r(\xi, \theta) = \tilde{W}(\xi \cos \theta, \xi \sin \theta)$ . Noting that  $\tilde{p}r(\xi, \theta) = \int_{-\infty}^{\infty} \langle q | U(\theta) \hat{\rho} U^\dagger(\theta) | q \rangle e^{-i\xi q} dq = \text{Tr} \hat{\rho} U^\dagger(\theta) e^{-i\xi q} U(\theta)$ , we can with a change of variables write  $\tilde{W}(u, v) = \text{Tr} \hat{\rho} e^{-iu\hat{q} - iv\hat{p}} = \text{Tr} \hat{\rho} D(\alpha)$ , where  $D(\alpha)$  is known as the displacement operator (this shall be more properly introduced in the next section).

Finally, by performing the inverse Fourier transform, we can find the explicit equation for a single mode Wigner function,  $W(q, p) = \frac{1}{\pi} \int_{-\infty}^{\infty} \langle q + x | \hat{\rho} | q - x \rangle e^{2ix \cdot p} dx$ . This expression can be generalised to  $N$  modes, to give [80]

$$W_{\hat{\rho}}(\mathbf{q}, \mathbf{p}) = \frac{1}{\pi^N} \int_{\mathbb{R}^N} \langle \mathbf{q} + \mathbf{x} | \hat{\rho} | \mathbf{q} - \mathbf{x} \rangle e^{2i\mathbf{x} \cdot \mathbf{p}} d^N \mathbf{x}, \quad \mathbf{q}, \mathbf{p} \in \mathbb{R}^N. \quad (3.9)$$

The Wigner function is normalised to 1, in that  $\int_{\mathbb{R}^{2N}} W_{\hat{\rho}}(\mathbf{q}, \mathbf{p}) d^N q d^N p = 1$ . This amounts to taking the trace of the density matrix, but it also fortifies the position of the Wigner function as bearing similarity to a probability distribution. We can similarly take partial traces by integrating only over the quadratures corresponding to the modes we wish to trace over.

An especially important characteristic of the Wigner function is the ability to define Wigner-type functions for operators (although there are some constraints related to boundedness) [70]. We can do this simply by replacing the density matrix in Equation (3.9) by the relevant operator; for e.g. an operator  $\hat{O}$ , we denote this function by  $W_{\hat{O}}(q, p)$ .

Using this, we obtain a simple formula for calculating expectation values,  $\text{Tr} \hat{\rho} \hat{O} = \pi^N \int_{\mathbb{R}^N} W_{\hat{\rho}}(\mathbf{q}, \mathbf{p}) W_{\hat{O}}(\mathbf{q}, \mathbf{p}) d^N q d^N p$ . This enables us to see how the Wigner distribution closely approximates a classical phase distribution in how it enables us to calculate moments.

The other quasiprobability distributions can be derived directly from the characteristic function. For consistency with convention we shall shift our notation and denote the (generally  $N$ -mode) characteristic function of a density matrix  $\hat{\rho}$  by  $\chi_{\hat{\rho}}(\boldsymbol{\xi})$ , where  $\boldsymbol{\xi}$  some  $2N$ -dimensional set of real parameters. Then we can define the  $s$ -ordered characteristic functions by,

$$\chi_{\hat{\rho}}^s(\boldsymbol{\xi}) = \chi_{\hat{\rho}}(\boldsymbol{\xi}) e^{s\|\boldsymbol{\xi}\|^2/2}. \quad (3.10)$$

These characteristic functions, when subjected to the inverse the Fourier transform, yield the  $s$ -ordered quasiprobability distributions  $W_{\hat{\rho}}^s(\mathbf{q}, \mathbf{p})$ . For  $s = 0$ , we retain the

### 3. CONTINUOUS VARIABLES

---

Wigner distribution, for  $s = 1$ , we have the  $P$ -distribution and for  $s = -1$ , we find the  $Q$ -distribution. This explains the earlier statement that the  $Q$ -distribution is obtained by the Wigner distribution by ‘smoothing’ it with a Gaussian function. Owing to the fact that the Fourier transform turns multiplications into convolutions, we see by performing the transform of  $\chi_\rho^{(-1)}(\boldsymbol{\xi})$ , we obtain a convolution of the Wigner function with a Gaussian.

The Wigner and  $Q$ -distributions exist for any density matrix, whereas the  $P$ -distribution does not, and in fact it may be ‘more singular than a Dirac delta’ [68]. The  $Q$ -distribution is also always positive, however, this does not make it useful in of itself because the corresponding operator density functions for calculating moments (i.e. the analogues to  $W_{\hat{O}}(\mathbf{q}, \mathbf{p})$ ) are generally very ill-behaved for the  $Q$ -distribution. Conversely, although the  $P$ -distribution is poorly behaved, when it does exist, we can almost always calculate moments easily because the density functions exist for the vast majority of operators.

We do not use the  $Q$ -distribution in this thesis, however both the Wigner and  $P$ -distributions are important, and the next section will consider the case of states with Gaussian Wigner functions. The  $P$ -distribution is important because it provides a notion of classicality in quantum optics. This can be seen by writing the density matrix explicitly with respect to the  $P$  distribution.

$$\hat{\rho} = \int_{\mathbb{R}^N} P(\alpha, \beta, \dots) |\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta| \otimes \dots \quad (3.11)$$

In effect, when the  $P$ -distribution is well-behaved, we have classical mixtures of coherent states. Given the aforementioned classical-like nature of coherent states, it is intuitively apparent that a classical mixture of classical-like states would be classical-like itself. Moreover, we can see that concepts like entanglement can’t feasibly arise with a well-behaved  $P$ -distribution due to the tensor product under the integral effectively implying the fulfilment of the separability criterion (Equation (2.7)). The connection of the  $P$ -distribution to semi-classical optics was first shown by Sudarshan [18].

### 3.3 Gaussian States

Gaussian functions are introduced early on in our learning of probability theory, often under the name of ‘normal distributions’. These functions appear endlessly throughout

the study of probability and statistics and it would be wise for any mathematician or physicist to be familiar with them. Though perhaps not as familiar a term, Gaussian *states* are analogously ubiquitous in the laboratories of quantum physicists: as we've already mentioned, coherent states, such as those from a laser; thermal states, as from a black body source, and even the vacuum state are all Gaussian. Importantly, Gaussian states are very closely related to Gaussian functions. A Gaussian state is defined as any state whose Wigner characteristic function and probability distribution are Gaussian functions on the quantum phase space. In the case of pure states, this property also coincides with a Gaussian wavefunction in the quadrature (position or momentum) basis [65].

A general multi-mode Gaussian function has the form

$$f(\mathbf{x}) = C \exp \left( -\frac{1}{2} \mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{b}^\top \mathbf{x} \right), \quad (3.12)$$

where  $\mathbf{x} = (x_1, x_2, \dots, x_N)^\top$ ,  $\mathbf{b} = (b_1, b_2, \dots, b_N)^\top$ , and  $\mathbf{A}$  is an  $N \times N$  positive-definite matrix.

A Gaussian state on the other hand, is defined as a function possessing Gaussian characteristic and Wigner functions. Conventionally, this is expressed as

$$\chi_{\hat{\rho}}(\boldsymbol{\xi}) = e^{-\frac{1}{4} \boldsymbol{\xi}^\top \boldsymbol{\Omega} \boldsymbol{\sigma} \boldsymbol{\Omega}^\top \boldsymbol{\xi} - i(\boldsymbol{\Omega} \mathbf{d})^\top \boldsymbol{\xi}}, \quad (3.13a)$$

$$W_{\hat{\rho}}(\mathbf{R}) = \frac{1}{\pi^N} \frac{1}{\sqrt{\det(\boldsymbol{\sigma})}} e^{-(\mathbf{R} - \mathbf{d})^\top \boldsymbol{\sigma}^{-1} (\mathbf{R} - \mathbf{d})}. \quad (3.13b)$$

For  $N$  modes,  $\boldsymbol{\Omega} = \bigoplus_1^n \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  is the symplectic form and the covariance matrix is expressed as  $\sigma_{ij} = \langle \hat{R}_i \hat{R}_j + \hat{R}_j \hat{R}_i \rangle - 2\langle \hat{R}_i \rangle \langle \hat{R}_j \rangle$ . The elements,  $\hat{R}_i$  are drawn from the vector of quadratures  $\hat{\mathbf{R}} = (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_N, \hat{p}_N)$  and we choose the convention such that the commutation relation reads as  $[\hat{R}_i, \hat{R}_j] = i\Omega_{ij}$ . It's important to note that another popular convention replaces  $i$  by  $\frac{i}{2}$  in the previous equation, in which case we also find factors of two appearing in the normalisation of the Wigner function and the covariance matrix. To complete the notation,  $\mathbf{d} = \langle \hat{\mathbf{R}} \rangle$  is known as the displacement vector and  $\mathbf{R}$  (sans caret), is the vector of quadrature eigenvalues with  $\boldsymbol{\xi}$  the Fourier transformed variables. The covariance matrix  $\boldsymbol{\sigma}$  will be the most important quantity for our purposes, as the displacements can be set to 0 by local unitary operations and thus the covariance matrix captures all the correlations of the system.

### 3. CONTINUOUS VARIABLES

---

As mentioned, for pure states, Gaussian states also have Gaussian wavefunctions and it is instructive to calculate the general sorts of operations which preserve Gaussianity on known Gaussian wavefunctions. The simplest Gaussian function is the single-mode vacuum state  $|0\rangle$  which is an eigenstate of the annihilation operator  $\hat{a}$  with eigenvalue 0. Expressing the annihilation operator in terms of quadratures,  $\hat{a} = \frac{1}{\sqrt{2}}(\hat{q} + i\hat{p})$ , we can easily evaluate the vacuum wavefunction expressed in the  $q$ -quadrature basis,  $\psi_0(q) = \langle q|0\rangle$ . We find

$$\hat{a}|0\rangle = \frac{1}{\sqrt{2}}(\hat{q} + i\hat{p}) \int dq|q\rangle \langle q|0\rangle = \int dq|q\rangle \left(q + \frac{\partial}{\partial q}\right) \langle q|0\rangle,$$

thus  $\left(q + \frac{\partial}{\partial q}\right) \psi_0(q) = 0$  and

$$\psi_0(q) = \frac{1}{\sqrt[4]{\pi}} e^{-\frac{q^2}{2}}. \quad (3.14)$$

In accordance with our expectations we also find the Wigner function of the vacuum state to be a Gaussian given by

$$W_{|0\rangle}(q, p) = \frac{1}{\pi} e^{-q^2 - p^2}. \quad (3.15)$$

This is easily checked using the expression for a single-mode pure state Wigner function, from Equation (3.9):  $W_{|\psi\rangle}(q, p) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{2ipx} \psi^*(q+x) \psi(q-x) dx$ .

It turns out that all possible single-mode Gaussian unitaries can be described by two operators. The first is the displacement operator defined by

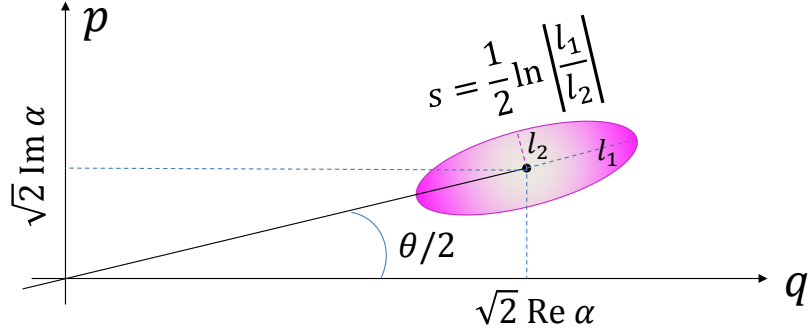
$$\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \bar{\alpha} \hat{a}}. \quad (3.16)$$

The effect of this operator on a vacuum state effectively transforms it into a coherent state so  $\hat{D}(\alpha)|0\rangle = |\alpha\rangle$ . In the phase-space picture, this effectively moves the minimum uncertainty Gaussian wave packet of the vacuum from the origin. The other necessary single mode transformation is the squeezing unitary,  $\hat{S}(\zeta)$ ,

$$\hat{S}(\zeta) = \exp \left[ \frac{1}{2} (\zeta \hat{a}^{\dagger 2} - \bar{\zeta} \hat{a}^2) \right], \quad \zeta = s e^{i\theta}, \quad (3.17)$$

Together, these can generate all single-mode Gaussian pure states,  $|\psi_G\rangle$ , by acting these operators sequentially:  $|\psi_G\rangle = D(\alpha)S(\zeta)|0\rangle$ . A depiction of the outcome of such a transformation can be seen in Figure 3.1.

In addition to these single mode transformations, we additionally only need two mode transformations to effectively describe all possible Gaussian unitaries – that is



**Figure 3.1: Cross-section of the Wigner function for a general pure Gaussian state**  $|\psi_{\alpha,s,\theta}\rangle = \hat{D}(\alpha)\hat{S}(se^{i\theta})|0\rangle$  of a single mode  $k$ , characterised by a complex displacement vector  $\alpha$ , a real squeezing degree  $s$  and a squeezing phase  $\theta$ . The Wigner function is given explicitly by the Gaussian form (3.13b). Figure from [4].

an  $N$ -mode unitary can be decomposed into products of one and two mode unitaries [65, 81]. We shall withhold the consideration of these unitaries until we have dissected the Gaussian formalism in some more detail.

### 3.3.1 Operations

Part of the true power of the covariance matrix formalism is that the effect of a large class of unitary operations on Gaussian states can be described by simple matrix multiplications with the covariance matrix. This is clearly desirable, since operations in the Fock space would be represented by infinite dimensional matrices.

In more detail, unitary transformations on a Hilbert space correspond to real *symplectic transformations* on the first and second moments as

$$\hat{\rho}' = \hat{U}^\dagger \hat{\rho} \hat{U} \rightarrow \begin{cases} \mathbf{d}' = \mathbf{S} \mathbf{d} \\ \boldsymbol{\sigma}' = \mathbf{S} \boldsymbol{\sigma} \mathbf{S}^\top \end{cases}, \quad (3.18)$$

where  $\mathbf{S}$  is a *symplectic matrix* which corresponds to the action of  $\hat{U}$  on the state  $\hat{\rho}$ .

This does not hold for all unitaries, but rather only for unitaries for which the generators are at most quadratic in the mode operators. This can be intuitively understood insofar as higher order terms in the mode-operators would affect higher than second-order moments, and since Gaussian states are fully determined by their second-order moments, Gaussianity would be compromised. These symplectic transformations form

### 3. CONTINUOUS VARIABLES

---

the building blocks of Gaussian state quantum information and we'll dedicate some space to properly understanding them. This section draws from [4, 82].

The group of real symplectic matrices is defined by the condition

$$\mathbf{S}\mathbf{\Omega}\mathbf{S}^\top = \mathbf{\Omega}, \quad (3.19)$$

where the matrix  $\mathbf{\Omega}$  defined below Equation (3.13b) is known as the symplectic form. This group is commonly denoted by  $\text{Sp}(2N, \mathbb{R})$  so that

$$\text{Sp}(2N, \mathbb{R}) = \left\{ \mathbf{S} \mid \mathbf{S}\mathbf{\Omega}\mathbf{S}^\top = \mathbf{\Omega} \right\}. \quad (3.20)$$

Symplectic matrices operating on  $N$ -mode systems are square  $(2N \times 2N)$ , non-singular matrices with unit determinant. The symplectic matrix can be decomposed as

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_{11} & \mathbf{s}_{12} & \cdots & \mathbf{s}_{1N} \\ \mathbf{s}_{21} & \mathbf{s}_{22} & & \\ \vdots & & \ddots & \\ \mathbf{s}_{N1} & & & \mathbf{s}_{NN} \end{pmatrix}, \quad (3.21)$$

where the  $2 \times 2$  sub-block  $\mathbf{s}_{mn}$  represents the transformation between the modes  $m$  and  $n$ . This relates back to the fact that Gaussian-preserving unitaries have exponents that are at most quadratic in the mode-operators, allowing at most pairwise mode-interactions.

Williamson showed that any symmetric positive-definite matrix can be put into a diagonal form via a symplectic transformation. An important use of this result, which amounts physically to a normal mode decomposition, is in finding the *symplectic eigenvalues* of an arbitrary Gaussian state. We can formalise this statement with Williamson's theorem [83]:

**Theorem 3.3.1.** *Let  $\boldsymbol{\sigma}$  be a  $2N \times 2N$  positive-definite matrix. Then there exists a unique  $\mathbf{S} \in \text{Sp}(2N, \mathbb{R})$  that diagonalises  $\boldsymbol{\sigma}$  such that*

$$\boldsymbol{\sigma} = \mathbf{S} \bigoplus_{k=1}^N \begin{pmatrix} \nu_k & 0 \\ 0 & \nu_k \end{pmatrix} \mathbf{S}^\top$$

We can write the  $N$  eigenvalues  $\nu_k$  in a diagonal matrix  $\boldsymbol{\nu} = \text{diag}(\nu_1, \dots, \nu_N)$ .  $\boldsymbol{\nu}$  is known as the *symplectic spectrum* of  $\boldsymbol{\sigma}$ . While a powerful result, the Williamson

theorem is not always the best route for finding the symplectic spectrum of a matrix  $\sigma$ , and it is usually more convenient to use the relation [64, 81]

$$\nu = \text{Eig}_+(i\Omega\sigma), \quad (3.22)$$

where  $\text{Eig}_+(\mathbf{A})$  denotes the diagonal matrix of *positive* (orthogonal) eigenvalues of the matrix  $\mathbf{A}$ . The  $N$  symplectic eigenvalues are thus determined by  $N$  invariants of the characteristic polynomial of the matrix  $|i\Omega\sigma|$  [84]. Knowing the symplectic spectrum of a given covariance matrix is very desirable for calculating informational measures. The symplectic eigenvalues are very closely related to the thermality and entropy of a state, a point which we will expand upon when looking at explicit covariance matrices for different types of states.

The next big step is to link known Gaussian unitary operations to their symplectic counterparts. However, since it is more conventional to express unitary operations in terms of mode operators rather than quadratures, we should first examine how to transform between each in the symplectic formalism. To do this, we only need to recall that  $\hat{a} = \frac{1}{\sqrt{2}}(\hat{q} + i\hat{p})$ . We also for convenience, rearrange the mode operators so that all the  $\hat{a}_k$  occur in order, followed by all the  $\hat{a}_k^\dagger$ . Thus, we can write,

$$\hat{\mathbf{M}} \equiv \begin{pmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_N \\ \hat{a}_1^\dagger \\ \vdots \\ \hat{a}_N^\dagger \end{pmatrix} = \mathbf{L}\mathbf{T} \begin{pmatrix} \hat{q}_1 \\ \hat{p}_1 \\ \vdots \\ \hat{q}_N \\ \hat{p}_N \end{pmatrix}, \quad (3.23)$$

where the transformation matrix  $\mathbf{T}$  has elements given by  $T_{jk} = \delta_{k,2j-1} + \delta_{k+2N,2i}$ , thus rearranging the quadratures, and the matrix  $\mathbf{L}$  can be written as,

$$\mathbf{L} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & i\mathbb{1} \\ \mathbb{1} & -i\mathbb{1} \end{pmatrix}. \quad (3.24)$$

and combines the quadratures to form the mode operators. Here  $\mathbb{1}$  is the  $N$ -dimensional identity matrix. We can convert any symplectic operation  $\mathbf{S}$  from the quadrature basis

### 3. CONTINUOUS VARIABLES

---

to the mode basis by  $\mathbf{S}_c = (\mathbf{L}\mathbf{T})\mathbf{S}(\mathbf{L}\mathbf{T})^\dagger$ . Equation 3.19, the defining equation of the symplectic matrices takes a new form,

$$\mathbf{K} = \mathbf{S}_c \mathbf{K} \mathbf{S}_c^\dagger, \quad \mathbf{K} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \quad (3.25)$$

In this basis it becomes very clear how Gaussian operations connect to linear optics. We take a generic Gaussian unitary  $\hat{U} = e^{-i\hat{H}}$ . Choosing to omit linear terms, the most general generator  $\hat{H}$  can be written as,

$$\hat{H} = A_{mn} \hat{a}_m^\dagger \hat{a}_n + B_{mn} \hat{a}_m^\dagger \hat{a}_n^\dagger + \bar{B}_{mn} \hat{a}_m \hat{a}_n + \bar{A}_{mn} \hat{a}_m \hat{a}_n^\dagger. \quad (3.26)$$

Acted on by such a unitary, the modes evolve by the equations,

$$\begin{aligned} \hat{U}^\dagger \hat{a}_k \hat{U} &= \sum_j \alpha_{jk} \hat{a}_j + \sum_j \beta_{jk} \hat{a}_j^\dagger, \\ \hat{U}^\dagger \hat{a}_k^\dagger \hat{U} &= \sum_j \bar{\alpha}_{jk} \hat{a}_j^\dagger + \sum_j \bar{\beta}_{jk} \hat{a}_j, \end{aligned} \quad (3.27)$$

or more compactly,

$$\hat{U}^\dagger \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{pmatrix} \hat{U} = \begin{pmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ \bar{\boldsymbol{\beta}} & \bar{\boldsymbol{\alpha}} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{a}}^\dagger \end{pmatrix}. \quad (3.28)$$

To ensure that the unitary operations are valid, we require the conditions,

$$\boldsymbol{\alpha} \boldsymbol{\alpha}^\dagger - \boldsymbol{\beta} \boldsymbol{\beta}^\dagger = \mathbb{1}, \quad (3.29a)$$

$$\boldsymbol{\alpha} \boldsymbol{\beta}^T = (\boldsymbol{\alpha} \boldsymbol{\beta}^T)^T, \quad (3.29b)$$

hold true. It turns out that this is equivalent as the defining conditions of Equation (3.25), and in fact we can indeed write for a quadratic Hamiltonian,

$$\mathbf{S}_c = \begin{pmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ \bar{\boldsymbol{\beta}} & \bar{\boldsymbol{\alpha}} \end{pmatrix} \quad (3.30)$$

Hence linear optics operations hold a central position in the theory of Gaussian states.



Ultimately, if we are given a generic quadratic Hamiltonian  $\hat{H}$ , then we can immediately deduce the corresponding symplectic matrix. This is done by noticing that we can write  $\hat{H} = \hat{\mathbf{M}}^\dagger \mathbf{H} \hat{\mathbf{M}}$ , where  $\mathbf{H}$  is the matrix of the form,

$$\mathbf{H} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \bar{\mathbf{B}} & \bar{\mathbf{A}} \end{pmatrix}, \quad (3.31)$$

with the entries of the submatrices  $\mathbf{A}$  and  $\mathbf{B}$  taken from Equation 3.26. The symplectic matrix can then be written as

$$\mathbf{S}_c = e^{-i\mathbf{K}\mathbf{H}}$$

where  $\mathbf{K}$  is as previously defined in Equation 3.25. We can subsequently revert back to the quadrature basis by inverting the transformation.

### 3.3.2 Single and two-mode operations

We've already clarified that all Gaussian operations can be decomposed into operations on one and two modes. Moreover, in this thesis we aim to deal with at-most two-mode systems, which is a necessary constraint due to the complexity of calculating correlations otherwise. It is conventional to use the quadrature basis in quantum information, and we shall do so henceforth unless otherwise stated.

We will now give examples of standard symplectic matrices  $\mathbf{S}$  associated to relevant linear optics transformations.

**Rotation.** A single-mode rotation by an angle  $\varphi/2$  in phase space, also known as phase shift, is the simplest example of a passive transformation (i.e. an energy conserving transformation). Its unitary form is

$$\hat{U} = \exp\left(i\varphi \hat{a}_k^\dagger \hat{a}_k\right)$$

for a mode  $k$ . This corresponds to a quadratic generator with matrix representation  $\mathbf{H} = -\frac{\varphi}{2}\mathbb{1}$ . In the quadrature basis, the symplectic transformation  $\mathbf{R}(\varphi)$  associated to

### 3. CONTINUOUS VARIABLES

---

a rotation can be obtained by

$$\begin{aligned}
\mathbf{R}(\varphi) &= \mathbf{T}^\top \mathbf{L}^\dagger e^{-i\mathbf{K}\mathbf{H}} \mathbf{L} \mathbf{T} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} e^{\frac{i\varphi}{2}} & 0 \\ 0 & e^{-\frac{i\varphi}{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \cos(\frac{\varphi}{2}) & -\sin(\frac{\varphi}{2}) \\ \sin(\frac{\varphi}{2}) & \cos(\frac{\varphi}{2}) \end{pmatrix}. \tag{3.32}
\end{aligned}$$

**Single-mode squeezing.** The single-mode squeezing operator is a prototypical active transformation (it increases the mean energy), described by the unitary operator  $\hat{S}_k(se^{i\theta})$  introduced in Equation (3.17). In this case, referring to Eq. (3.31), we have  $\mathbf{A} = 0$  and  $\mathbf{B} = is e^{i\theta}$ . Adopting the same procedure as before, we obtain the symplectic representation of squeezing,

$$\mathbf{S}^{(1)}(s, \theta) = \begin{pmatrix} \cosh(s) + \cos(\theta) \sinh(s) & \sin(\theta) \sinh(s) \\ \sin(\theta) \sinh(s) & \cosh(s) - \cos(\theta) \sinh(s) \end{pmatrix}, \tag{3.33}$$

which reduces to  $\mathbf{S}^{(1)}(s, 0) = \text{diag}(e^s, e^{-s})$  for  $\theta = 0$ . In the latter case, this operation (for  $s > 0$ ) squeezes the momentum, reducing its variance exponentially, while correspondingly enlarging the one on position. The complementary case  $\theta = \pi/2$  amounts to a squeeze of the position quadrature and a corresponding increase on the variance of the momentum quadrature.

We can now write the phase space representation of the most general pure single-mode Gaussian state  $|\psi_G\rangle = D(\alpha)S(\zeta)|0\rangle$ . We just need to apply the operation in Equation (3.33) to the vacuum state, followed by a displacement. Recall that the vacuum has covariance matrix equal to the identity, and vanishing first moments. The first and second moments of a general pure single-mode Gaussian state  $|\psi_G\rangle$  are then given by

$$\begin{aligned}
\mathbf{d} &= \sqrt{2} \begin{pmatrix} \text{Re}(\alpha) \\ \text{Im}(\alpha) \end{pmatrix}, \\
\boldsymbol{\sigma} &= \mathbf{S}^{(1)}(s, \theta) \mathbf{1} \mathbf{S}^{(1)\top}(s, \theta) \\
&= \begin{pmatrix} \cosh(2s) + \cos(\theta) \sinh(2s) & \sin(\theta) \sinh(2s) \\ \sin(\theta) \sinh(2s) & \cosh(2s) - \cos(\theta) \sinh(2s) \end{pmatrix}. \tag{3.34}
\end{aligned}$$

Inserting this expression into Equation (3.13b) one obtains the Wigner function whose cross-section has been depicted in Figure 3.1.

**Beam splitter.** Another common unitary operation is the ideal (phase-free) *beam splitter*, whose action  $\hat{B}_{i,j}$  on a pair of modes  $i$  and  $j$  is defined as

$$\hat{B}_{i,j}(\theta) : \begin{cases} \hat{a}_i \rightarrow \hat{a}_i \cos \theta + \hat{a}_j \sin \theta \\ \hat{a}_j \rightarrow \hat{a}_i \sin \theta - \hat{a}_j \cos \theta \end{cases}. \quad (3.35)$$

A beam splitter with transmissivity  $\tau$  is a passive transformation corresponding to a rotation of  $\theta = \arccos \sqrt{\tau}$  in phase space; in particular,  $\theta = \pi/4$  corresponds to a balanced 50:50 beam splitter,  $\tau = 1/2$ . Applying the machinery introduced above, one finds that the beam splitter is described by a symplectic transformation

$$\mathbf{B}_{i,j}(\tau) = \begin{pmatrix} \sqrt{\tau} & 0 & \sqrt{1-\tau} & 0 \\ 0 & \sqrt{\tau} & 0 & \sqrt{1-\tau} \\ \sqrt{1-\tau} & 0 & -\sqrt{\tau} & 0 \\ 0 & \sqrt{1-\tau} & 0 & -\sqrt{\tau} \end{pmatrix}. \quad (3.36)$$

**Two-mode squeezing.** We close this gallery with the two-mode squeezing operation, an active transformation which models the physics of optical parametric amplifiers (see e.g. [85]) and is routinely employed to create CV entanglement. Acting on the pair of modes  $i$  and  $j$  via the unitary

$$\hat{U}_{i,j}(r) = \exp[r(\hat{a}_i^\dagger \hat{a}_j^\dagger - \hat{a}_i \hat{a}_j)], \quad (3.37)$$

it corresponds to the symplectic transformation

$$\mathbf{S}_{i,j}^{(2)}(r) = \begin{pmatrix} \cosh r & 0 & \sinh r & 0 \\ 0 & \cosh r & 0 & -\sinh r \\ \sinh r & 0 & \cosh r & 0 \\ 0 & -\sinh r & 0 & \cosh r \end{pmatrix}. \quad (3.38)$$

### 3.3.3 Partial tracing and loss

This section will be concluded by considering some non-unitary operations. Partial tracing in the covariance matrix formalism is actually a remarkably easy process. We

### 3. CONTINUOUS VARIABLES

---

are aided in this task by writing the displacement vector and covariance matrix in terms of two-dimensional subblocks

$$\mathbf{d}_{1,\dots,N} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_N \end{pmatrix}, \quad \boldsymbol{\sigma}_{1,\dots,N} = \begin{pmatrix} \boldsymbol{\sigma}_1 & \boldsymbol{\varepsilon}_{1,2} & \cdots & \boldsymbol{\varepsilon}_{1,N} \\ \boldsymbol{\varepsilon}_{1,2}^\top & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{\varepsilon}_{N-1,N} \\ \boldsymbol{\varepsilon}_{1,N}^\top & \cdots & \boldsymbol{\varepsilon}_{N-1,N}^\top & \boldsymbol{\sigma}_N \end{pmatrix}. \quad (3.39)$$

If we then wish to trace out, without loss of generality, the modes  $3, \dots, N$ , then all we need to do is erase the corresponding entries of the covariance matrix to find the reduced matrix of modes 1 and 2 to be

$$\mathbf{d}_{1,2} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}, \quad \boldsymbol{\sigma}_{1,2} = \begin{pmatrix} \boldsymbol{\sigma}_1 & \boldsymbol{\varepsilon}_{1,2} \\ \boldsymbol{\varepsilon}_{1,2}^\top & \boldsymbol{\sigma}_2 \end{pmatrix}. \quad (3.40)$$

This turns out to be especially useful when considering lossy channels. A lossy channel can be characterised by a beamsplitter transformation, followed by a tracing out of the auxilliary mode [78]. This can also be used to model an inefficient detector by considering a perfect detector behind a lossy beam splitter [86].

#### 3.3.4 Gaussian Measurements

In this section we consider in general a POVM. A Gaussian measurement is one which maps Gaussian states to Gaussian states. Any such measurement can be realised by appending ancillae to the system, performing symplectic operations on the composite state of system and ancillae and then measuring the quadratures, e.g. by homodyne detection [87]. For a bipartite Gaussian state  $\hat{\rho}_{AB}$  with a measurement on Bob, where Bob holds  $N_B$  modes, the most general Gaussian POVM is described by

$$\hat{\Pi}_B(\boldsymbol{\eta}) = \pi^{-N_B} \left[ \prod_{j=1}^{N_B} \hat{D}_{B_j}(\eta_j) \right] \Lambda_B^{\hat{\Pi}} \left[ \prod_{j=1}^{N_B} \hat{D}_{B_j}^\dagger(\eta_j) \right], \quad (3.41)$$

where

$$\hat{D}_B(\eta_j) = \exp(\eta_j \hat{b}_j^\dagger - \bar{\eta}_j \hat{b}_j) \quad (3.42)$$

is the displacement operator of Equation (3.16),  $\hat{b}_j$  is the annihilation operator of the  $j$ -th mode of the subsystem  $B$ ,  $\pi^{-N_B} \int \hat{\Pi}_B(\boldsymbol{\eta}) d^{2N_B} \boldsymbol{\eta} = 1$ , and  $\Lambda_B^{\hat{\Pi}}$  is the density matrix of

a (generally mixed)  $N_B$ -mode Gaussian state with covariance matrix  $\mathbf{\Gamma}_B^{\hat{\Pi}}$  which denotes the so-called seed of the measurement. The conditional state  $\hat{\rho}_{A|\boldsymbol{\eta}}$  of subsystem  $A$  after the measurement  $\hat{\Pi}_B(\boldsymbol{\eta})$  has been performed on  $B$  has a covariance matrix  $\tilde{\boldsymbol{\sigma}}_A^{\hat{\Pi}}$  independent of the outcome  $\boldsymbol{\eta}$  and given by the Schur complement [88, 89, 90]

$$\tilde{\boldsymbol{\sigma}}_A^{\hat{\Pi}} = \boldsymbol{\sigma}_A - \boldsymbol{\varepsilon}_{AB}(\boldsymbol{\sigma}_B + \mathbf{\Gamma}_B^{\hat{\Pi}})^{-1}\boldsymbol{\varepsilon}_{AB}^T, \quad (3.43)$$

where the original bipartite covariance matrix  $\boldsymbol{\sigma}_{AB}$  of the  $N$ -mode state  $\hat{\rho}_{AB}$  has been written in block form as in Equation (3.40).

### 3.4 Entropy and Correlations

Now that we have considered two-mode operations, we are adequately equipped to unveil the theory of bipartite correlations. This shall be done in a similar fashion to the first chapter, in terms of entropic measures, but with a necessarily greater emphasis on rigour and detail. It is useful, however, to begin the section with consideration of one of the archetypical quantum-correlated states, that first introduced by Einstein, Podolsky and Rosen.

#### 3.4.1 EPR Correlations

Gaussian states allow us to come close to the so-called *EPR state* of the famed paper of Einstein, Podolsky and Rosen [91] which first identified some of the seemingly paradoxical features of entanglement. This state is represented by

$$|\psi_{EPR}\rangle_{AB} = \delta(\hat{q}_A - \hat{q}_B)\delta(\hat{p}_A - \hat{p}_B) \quad (3.44)$$

This state is technically unphysical; it contains infinite energy and thus cannot be normalised. However, it can be asymptotically approached by certain states, the most important of which is the Gaussian *two-mode squeezed state* (TMSS). Acting the two-mode squeezing operator on the vacuum state yields the output covariance matrix

$$\sigma_{TMSS}(r) = \begin{pmatrix} \cosh(2r) & 0 & \sinh(2r) & 0 \\ 0 & \cosh(2r) & 0 & -\sinh(2r) \\ \sinh(2r) & 0 & \cosh(2r) & 0 \\ 0 & -\sinh(2r) & 0 & \cosh(2r) \end{pmatrix} \quad (3.45)$$

### 3. CONTINUOUS VARIABLES

---

#### 3.4.2 Entropy

We identified earlier that the symplectic eigenvalues are closely related to the thermal-ity, thus mixedness, thus entropy of Gaussian states. We can in fact denote the purity of a Gaussian state by writing,

$$\mu_{\hat{\rho}} = \text{Tr } \hat{\rho}^2 = \frac{1}{\sqrt{\det \boldsymbol{\sigma}}}, \quad (3.46)$$

which can easily be deduced from the Wigner function and the ‘overlap formula’. Noting that  $\det \boldsymbol{\sigma} = \Pi_k \nu_k^2$ , we see this is purely a function of the symplectic eigenvalues. Only when every symplectic eigenvalue is equal to 1 do we have a pure state.

The von Neumann entropy for an  $N$ -dimensional Gaussian state with covariance matrix  $\boldsymbol{\sigma}$  and symplectic spectrum  $\boldsymbol{\nu}$  is given by

$$S(\boldsymbol{\sigma}) = \sum_{i=1}^N \left( \frac{\nu_i + 1}{2} \log \frac{\nu_i + 1}{2} - \frac{\nu_i - 1}{2} \log \frac{\nu_i - 1}{2} \right). \quad (3.47)$$

Thus again, the symplectic spectrum is of paramount importance. Similarly, the Rényi- $\alpha$  entropy is given by

$$S_\alpha(\boldsymbol{\sigma}) = \frac{1}{1 - \alpha} \sum_{i=1}^N \log g_\alpha(\nu_i) \quad (3.48)$$

where  $g_\alpha(x) = 2^\alpha / [(x + 1)^\alpha - (x - 1)^\alpha]$ . In the special case of the Rényi-2 entropy we readily find  $S_2(\hat{\rho}) = -\log \text{Tr } \hat{\rho}^2 = \frac{1}{2} \log \det \boldsymbol{\sigma}$ . Interestingly, this is actually the Shannon entropy of the Wigner distribution modulo an additive constant [60],

$$H(W_{\hat{\rho}}(\boldsymbol{\xi})) = - \int W_{\hat{\rho}}(\boldsymbol{\xi}) \log \{W_{\hat{\rho}}(\boldsymbol{\xi})\} d^{2N} \boldsymbol{\xi} \quad (3.49)$$

$$= \mathcal{S}_2(\hat{\rho}) + N(1 + \log \pi). \quad (3.50)$$

As previously mentioned, the Rényi-2 entropy satisfies the strong subadditivity property. Using the explicit expression for Gaussian states this implies

$$\frac{1}{2} \log \left( \frac{\det \boldsymbol{\sigma}_{AB} \det \boldsymbol{\sigma}_{BC}}{\det \boldsymbol{\sigma}_{ABC} \det \boldsymbol{\sigma}_B} \right) \geq 0. \quad (3.51)$$

This result follows from applying the Hadamard-Fischer inequality [92] to the covariance matrix  $\boldsymbol{\sigma}_{ABC}$ . This inequality states, for an  $n$ -dimensional positive Hermitian matrix

$\mathbf{A}$ , and any given set of indices  $\alpha = (\alpha_1, \alpha_2, \dots)$  and  $\beta = (\beta_1, \beta_2, \dots)$  where  $\alpha_i, \beta_i \in (1, \dots, n)$ , that

$$\det(\mathbf{A}_{\alpha \cup \beta}) \det(\mathbf{A}_{\alpha \cap \beta}) \leq \det(\mathbf{A}_\alpha) \det(\mathbf{A}_\beta) \quad (3.52)$$

where by  $\mathbf{A}_\alpha$  we mean the principal submatrix where the rows and columns are given by the set of indices  $\alpha$  and similarly for  $\beta$  and their union and intersection. By recognising that the covariance matrix is indeed positive definite, and identifying  $\alpha$  with the indices for modes  $AB$  and  $\beta$  with  $BC$  this immediately reduces to Equation (3.51). We thus can, and will, develop correlation measures for both types of entropy and the next several sections will be in large part book-keeping.

#### 3.4.3 Mutual Information

Recalling Equation (2.23), we write down

$$\mathcal{T}(A : B) = \mathcal{S}(\hat{\rho}_A) + \mathcal{S}(\hat{\rho}_B) - \mathcal{S}(\hat{\rho}_{AB}). \quad (3.53)$$

For the Rényi-2 entropy, this takes the remarkably simple form of

$$\begin{aligned} \mathcal{T}_2(\hat{\rho}_{A:B}) &= \mathcal{S}_2(\hat{\rho}_A) + \mathcal{S}_2(\hat{\rho}_B) - \mathcal{S}_2(\hat{\rho}_{AB}) \\ &= \frac{1}{2} \log \left( \frac{\det \boldsymbol{\sigma}_A \det \boldsymbol{\sigma}_B}{\det \boldsymbol{\sigma}_{AB}} \right). \end{aligned}$$

For the von Neumann entropy, it is not so informative to write down the general form of the mutual information. We will thus consider only the expression for two-modes. This requires us to first write the covariance matrix in *standard form*,

$$\begin{aligned} \boldsymbol{\sigma}_{AB} &= \begin{pmatrix} a & 0 & c_+ & 0 \\ 0 & a & 0 & c_- \\ c_+ & 0 & b & 0 \\ 0 & c_- & 0 & b \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\sigma}_A & \boldsymbol{\varepsilon}_{AB} \\ \boldsymbol{\varepsilon}_{AB}^T & \boldsymbol{\sigma}_B \end{pmatrix}, \end{aligned}$$

### 3. CONTINUOUS VARIABLES

---

where  $\alpha$ ,  $\beta$  and  $\gamma$  are  $2 \times 2$  sub-blocks from which we define *symplectic invariants*, quantities which do not change under the action of symplectic operations,

$$A = \det \sigma_A, \quad B = \det \sigma_B, \quad C = \det \epsilon_{AB}, \quad D = \det \sigma_{AB}. \quad (3.54)$$

The quantities  $A$ ,  $B$  and  $C$  are *local invariants*, in that they do not change under local symplectic operations but may under a global one, and  $D$  is globally invariant. We can define another global invariant, the *seralian*  $\Delta$ , as the sum of all the  $2 \times 2$  blocks, so for two modes  $\Delta = A + B + 2C$ . This allows a simple algebraic formula for the two symplectic eigenvalues

$$2\nu_{\pm}^2 = \Delta \pm \sqrt{\Delta^2 - 4D} \quad (3.55)$$

where  $\nu_-$  and  $\nu_+$  refer to the smaller and larger eigenvalues respectively. This finally allows us to write the von Neumann mutual information of two modes as

$$\mathcal{T}(\sigma_{AB}) = f(\sqrt{A}) + f(\sqrt{B}) - f(\nu_-) - f(\nu_+) \quad (3.56)$$

where  $f(x) = \frac{x+1}{2} \log \frac{x+1}{2} - \frac{x-1}{2} \log \frac{x-1}{2}$ .

#### 3.4.4 Classical and quantum correlations

For two modes it is possible to divide the mutual information into classical and quantum correlations, as in the sense of discord. In principle it should be possible for larger numbers of modes, but the problem is one of tractability due to the optimisation problem involved.

Recalling that classical correlations are defined with the optimisation problem

$$\mathcal{J}(\hat{\rho}_{A|B}) = S(\hat{\rho}_A) - \inf_{\{\Pi_i\}} S_{\{\Pi_i\}}(\hat{\rho}_{A|B}) \quad (3.57)$$

we wish to perform this optimisation over all Gaussian POVMs. It may seem that the optimal measurement may be non-Gaussian, however, for a large class of states (the type of which we tend to use in this thesis), it was shown that indeed the optimal measurement is a Gaussian one [93].

Considering two modes, with a measurement on Bob's mode, Alice's conditional state  $\tilde{\sigma}_A^{\hat{\Pi}}$  is independent on Bob's measurement outcome and is given by Equation 3.43.



The explicit expressions for the von Neumann and Rényi-2 classical correlations are, respectively

$$\mathcal{J}(\hat{\rho}_{A|B}) = \sup_{\Gamma_B^{\hat{\Pi}}} \left( f(\sqrt{A}) - f(\sqrt{\det \tilde{\sigma}_A^{\hat{\Pi}}}) \right)$$

$$\mathcal{J}_2(\hat{\rho}_{A|B}) = \sup_{\Gamma_B^{\hat{\Pi}}} \frac{1}{2} \left( \frac{\det \sigma_A}{\det \tilde{\sigma}_A^{\hat{\Pi}}} \right).$$

The minimisation of  $\det \tilde{\sigma}_A^{\hat{\Pi}}$  over all Gaussian POVMs is explicitly solved in [94]. For a given state, the only variable in the optimisation is the seed of the POVM  $\Gamma_B^{\hat{\Pi}}$ . For a two-mode state with a measurement on one mode, we can choose this to be the most general pure state covariance matrix: a rotated, squeezed matrix, which can be obtained from the vacuum by using the symplectic matrices Equations (3.32) and (3.33), giving  $\Gamma_B^{\hat{\Pi}} = \mathbf{R}(\varphi) \text{diag}(\lambda, \frac{1}{\lambda}) \mathbf{R}^T(\varphi)$ , where  $\lambda = e^{-r}$ , captures the squeezing of the state. Thus, for a matrix in standard form, we get the determinant

$$\det \tilde{\sigma}_A^{\Pi_{\lambda, \varphi}} = \frac{2a^2(b+\lambda)(1+b\lambda) - a(c_+^2 + c_-^2)(2b\lambda + \lambda^2 + 1) + 2c_+^2 c_-^2 \lambda + a(c_+^2 - c_-^2)(\lambda^2 - 1) \cos(2\varphi)}{2(b+\lambda)(1+b\lambda)}. \quad (3.58)$$

Optimising over  $\lambda$  and  $\varphi$  we can get

$$\inf_{\lambda, \varphi} \det \tilde{\sigma}_A^{\Pi_{\lambda, \varphi}} = \quad (3.59)$$

$$\begin{cases} a \left( a - \frac{c_+^2}{b} \right), & \text{if } (ab^2 c_-^2 - c_+^2(a + bc_-^2))(ab^2 c_+^2 - c_-^2(a + bc_+^2)) < 0; \\ \frac{2|c_- c_+| \sqrt{(a(b^2 - 1) - bc_-^2)(a(b^2 - 1) - bc_+^2)} + (a(b^2 - 1) - bc_-^2)(a(b^2 - 1) - bc_+^2) + c_-^2 c_+^2}{(b^2 - 1)^2}, & \text{otherwise.} \end{cases}$$

Inserting this into the equations for  $\mathcal{J}(\hat{\rho}_{A|B})$  and  $\mathcal{J}_2(\hat{\rho}_{A|B})$  gives us the explicit form of classical correlations. Then we can of course find the discord by recalling that

$$\mathcal{D}(\hat{\rho}_{A|B}) = \mathcal{T}(\hat{\rho}_{AB}) - \mathcal{J}(\hat{\rho}_{A|B}),$$

$$\mathcal{D}_2(\hat{\rho}_{A|B}) = \mathcal{T}_2(\hat{\rho}_{AB}) - \mathcal{J}_2(\hat{\rho}_{A|B}).$$

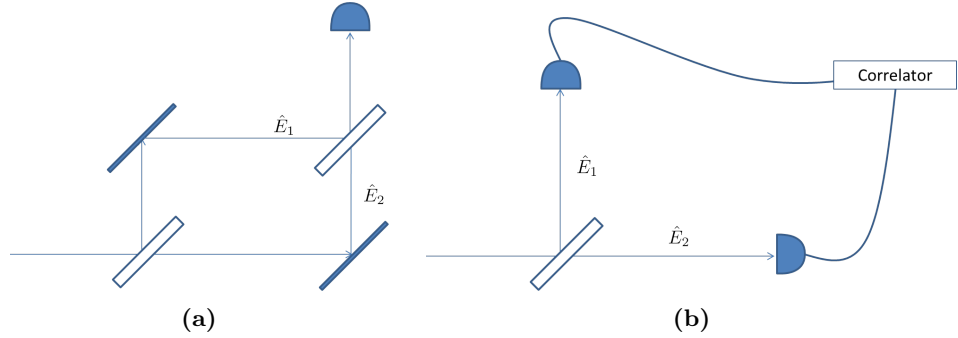
We can of course also consider a conditional measurement on Alice instead of Bob by simply switching the indices.

## Part II

# 4

## Intensity Interferometry

Intensity interferometry has been a source of persistent contention in physics since the 1950s, with the discovery of the Hanbury Brown and Twiss effect [27, 95, 96, 97, 98]. Interferometric experiments form the foundational backbone for enormous swathes of modern physics and amongst them are some of the most famous experiments in the field, from the Michelson and Morley null result for the existence of the luminiferous aether to Young’s two slit experiment for proving the wave nature of light (and subsequently the wave nature of quantum ‘particles’). Intensity interferometry, though a somewhat



**Figure 4.1: (a) Mach-Zehnder and (b) intensity interferometer.** The solid rectangles represent mirrors, and the empty ones represent balanced beam splitters. Despite their depiction, the optical paths in each arm of the detectors need not be the same length. The primary difference between the two is that the Mach-Zehnder interferometer measures the intensity after superposing  $\hat{E}_1$  and  $\hat{E}_2$ , whereas the intensity interferometer measures the intensities separately and then correlates the outputs.

less famous interferometric variant, is not so different in this respect as it was part of

## 4. INTENSITY INTERFEROMETRY

---

the driving force behind the work of early quantum opticians such as Glauber [77] and Sudarshan [18]. In Figure 4.1 we depict, the familiar Mach-Zehnder interferometer and an intensity interferometer; the fundamental difference lies in that the intensity interferometer aims to measure a second order correlation

$$g^{(2)}(\mathbf{r}_1, t_1; \mathbf{r}_2, t_2) = \frac{\langle \hat{E}^\dagger(\mathbf{r}_1, t_1) \hat{E}(\mathbf{r}_1, t_1) \hat{E}^\dagger(\mathbf{r}_2, t_2) \hat{E}(\mathbf{r}_2, t_2) \rangle}{\langle \hat{E}^\dagger(\mathbf{r}_1, t_1) \hat{E}(\mathbf{r}_1, t_1) \rangle \langle \hat{E}^\dagger(\mathbf{r}_2, t_2) \hat{E}(\mathbf{r}_2, t_2) \rangle},$$

while the standard interferometer only the first order correlation

$$g^{(1)}(\mathbf{r}_1, t_1; \mathbf{r}_2, t_2) = \frac{\langle \hat{E}^\dagger(\mathbf{r}_1, t_1) \hat{E}(\mathbf{r}_2, t_2) \rangle}{(\langle \hat{E}^\dagger(\mathbf{r}_1, t_1) \hat{E}(\mathbf{r}_1, t_1) \rangle \langle \hat{E}^\dagger(\mathbf{r}_2, t_2) \hat{E}(\mathbf{r}_2, t_2) \rangle)^{1/2}}.$$

Here  $\hat{E}$  is effectively equivalent to the usual annihilation operator  $\hat{a}$ , and  $[\hat{E}(\mathbf{r}_1, t_1), \hat{E}^\dagger(\mathbf{r}_2, t_2)] = \delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(t_1 - t_2)$ . There is some loss of phase information in the intensity interferometer; accordingly, in situations where turbulence of the medium, and thus phase distortions, are significant, intensity interferometry setups can be useful for imaging and detection schemes [99].

Recently, there has been a resurgence in intensity interferometry schemes due to the discovery of a number of powerful practical applications. Amongst, these are ghost imaging [100, 101, 102], quantum illumination [103, 104], and a number of variant schemes for popular quantum tasks such as QKD [105]. However, with this resurgence, has also rearisen a classic controversy of the nature of ‘quantumness’ present in intensity interferometry [26, 106, 107, 108]. In this chapter, the history of this controversy will be gradually unravelled, beginning with the Hanbury Brown and Twiss experiment itself, followed by ghost imaging, which is actually where the work on this thesis began, and finishing with quantum illumination which allowed some of the observations of this work on intensity interferometry to be experimentally tested (albeit in a limited way). Rather than tackling the problem from the perspective of whether a classical model (i.e. semi-classical optics) is permitted, we’ll consider more abstract correlations, such as discord.

### 4.1 Hanbury Brown and Twiss

As mentioned, the Hanbury Brown and Twiss experiment (Figure 4.2) is of great importance to the history of quantum optics. It caused ripples in the physics community

at the time, not because it was the first intensity interferometry experiment, but because it was the first intensity interferometry experiment *in the optical regime*. Such experiments had been performed in the microwave regime and radio astronomy [109], where there are large numbers of photons per mode and it is intuitively expected that wave-like behaviour arises, and spatial coherence of intensity between distant points is plausible.

However, when the situation is such that the light is sufficiently scarce that detection events are caused by individual photons, it becomes harder to conceptually understand how a correlation can still exist. The 1956 HBT paper [27] involved measuring the correlations between optical photons produced by a mercury arc lamp, which led to some publications [97, 98] questioning the veracity of the experimental results. However, Purcell [98] readily accepted and accounted for this phenomenon with a simple model. Moreover, Hanbury Brown and Twiss settled the discussion with a pair of detailed papers which provided both a classical and quantum account of the experiment [95, 96].

Despite this settlement, the Hanbury Brown and Twiss experiment had a resounding effect. It effectively led to the question, when are Maxwell's equations still a valid description of the light field? The answer to this forms the basis of semi-classical optics. More importantly, the answer to the question of when they are *not* valid, yielded quantum optics in the form it exists today.

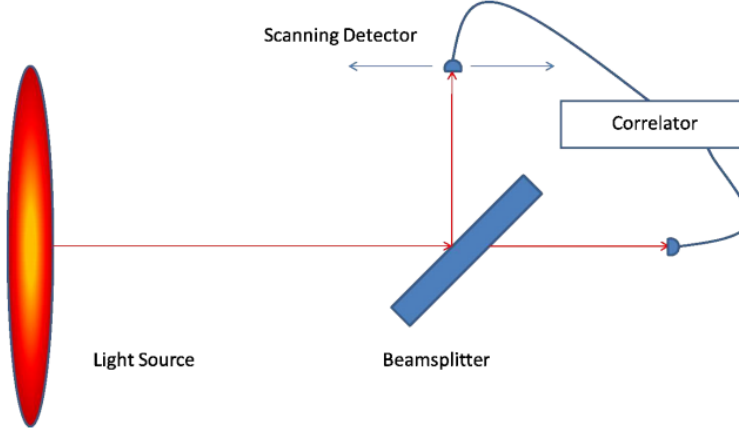
This section shall primarily serve as motivation for the coming section on ghost imaging, where we will provide a more complete account of intensity interferometry controversy which we will gloss over in this section. Here, it shall suffice to calculate the classical and quantum correlations present in the HBT scheme (in the sense of discord). This will reveal that a fresh perspective of what quantumness is may partially resolve the debate.

### 4.1.1 Covariance Matrix

To calculate the discord, we need to derive the covariance matrix in standard form. We will consider point-to-point spatial correlations, and a two-mode description is thus sufficient. Recalling the standard form, we wish to deduce the entries  $a$ ,  $b$ ,  $c$  and  $d$ .

#### 4. INTENSITY INTERFEROMETRY

---



**Figure 4.2: Diagrammatic representation of a Hanbury Brown Twiss setup.**

Light from a large source propagates into the far field whereupon it is split on a balanced beam-splitter. When one of the detectors is scanned in the transverse plane, the normalised intensity correlations scale as  $\frac{2J(\frac{k_0 r}{z}|\mathbf{x}-\mathbf{x}'|)}{\frac{k_0 r}{z}|\mathbf{x}-\mathbf{x}'|}$ , where  $J$  is a Bessel function of the first kind and first order and  $\mathbf{x} - \mathbf{x}'$  is the difference in the transverse positions of the detectors in each arm. Here  $r$  refers to the radius of the disc and  $z$  the distance of the detectors from disk

$$\begin{pmatrix} a & 0 & c & 0 \\ 0 & a & 0 & d \\ c & 0 & b & 0 \\ 0 & d & 0 & b \end{pmatrix}. \quad (4.1)$$

The entries are easily deduced with some simplifying assumptions: firstly, we choose a large disk-like monochromatic source, emitting spatially incoherent thermal light. From this, we take each mode  $\hat{a}(\mathbf{k})$  (where  $\mathbf{k}$  is the transverse component of the wave-vector) to be uncorrelated with every other mode in the source plane, and we consider all modes to be zero-mean modes with identical photon-number expectation value  $\bar{n}$ , such that  $\langle \hat{a}(\mathbf{k})^\dagger \hat{a}(\mathbf{k}') \rangle = \bar{n} \delta(\mathbf{k} - \mathbf{k}')$ , these are common assumptions for intensity interferometric experiments with thermal light [107] [110]. This implies that in the source plane  $\langle \hat{a}(\mathbf{x}_2)^\dagger \hat{a}(\mathbf{x}_1) \rangle = \int e^{i\mathbf{k} \cdot \mathbf{x}_1} e^{i\mathbf{k}' \cdot \mathbf{x}_2} \langle \hat{a}^\dagger(\mathbf{k}') \hat{a}(\mathbf{k}) \rangle d\mathbf{k} d\mathbf{k}' = \bar{n} \delta(\mathbf{x}_1 - \mathbf{x}_2)$ , where  $\mathbf{x}_i$  represent transverse position vectors. This is consistent with the intuitive picture of a large thermal source (such as a star) as being composed of large numbers of separate emit-

ters, thus entailing no spatial correlations at source. We first calculate the covariance matrix in the mode basis after which we can rotate to the quadrature basis. Taking a balanced beam splitter, we can relate the output modes  $\hat{b}_{1,2}$ , and the input modes of the beam splitter by

$$\hat{b}_{1,2}(\mathbf{x}) = \frac{1}{\sqrt{2}}(\hat{a}(\mathbf{x}) \pm \hat{a}_{vac}(\mathbf{x})). \quad (4.2)$$

Here,  $\hat{a}(\mathbf{x})$  represents the input from the distant planar source at the beam splitter plane and  $\hat{a}_{vac}(\mathbf{x})$  represents the vacuum mode which is combined with  $\hat{a}(\mathbf{x})$  by the beam splitter. Since we are interested in the spatial correlation, these are indexed by the transverse vector  $\mathbf{x}$ .

We consider the free-space propagation of light from the source to the beam splitter plane. Although we consider the far field regime, we do not assume emission from an infinite plane, and hence we are not making the approximation that we have only point-to-point (delta function) correlations as is often made in the ‘paraxial’ regime – where the region of interest is small in comparison to the emitter and is additionally located close to its optical axis.

We *desire* spatial information (which is why we don’t assume an infinite emitter), and thus require a setup where the spatial resolution is high in comparison to the coherence area of the light. This is feasible, for example, in the case of measuring distant stars, where the coherence length is of the order of metres [68]. Historically, the HBT scheme was indeed advertised as a stellar interferometer with which the diameter of stars could be measured.

The modes in the observation plane can be calculated with  $\hat{a}(\mathbf{x}) \propto \int_A g(\mathbf{x}_1, \mathbf{x}) \hat{a}(\mathbf{x}_1) d\mathbf{x}_1$ , where  $A$  is the source area. The parameter  $\mathbf{x}_1$ , refers to the transverse vector in the source plane, and  $\hat{a}(\mathbf{x}_1)$  the corresponding mode. In the far field, we use the Fraunhofer propagator  $g(\mathbf{x}_1, \mathbf{x}) = e^{-i\frac{k_0}{z}(\mathbf{x}_1 \cdot \mathbf{x})}$  where  $k_0$  is the mean wave-vector. As an example, we can then calculate the correlation function  $\langle \hat{a}^\dagger(\mathbf{x}) \hat{a}(\mathbf{x}') \rangle$ . Note that the indexed parameters  $\hat{a}(\mathbf{x}_1)$ ,  $\hat{a}(\mathbf{x}_2)$  refer to the modes in the source plane with transverse coordinates  $\mathbf{x}_1$  and  $\mathbf{x}_2$  respectively, whereas the position vectors without indices  $\mathbf{x}$  and  $\mathbf{x}'$  are used to describe the propagated operators  $\hat{a}(\mathbf{x})$  and  $\hat{a}(\mathbf{x}')$ .

#### 4. INTENSITY INTERFEROMETRY

---

$$\begin{aligned}
\langle \hat{a}(\mathbf{x})^\dagger \hat{a}(\mathbf{x}') \rangle &\propto \int_A \int_A g(\mathbf{x}_1, \mathbf{x}) g^*(\mathbf{x}_2, \mathbf{x}') \langle \hat{a}(\mathbf{x}_1) \hat{a}(\mathbf{x}_2) \rangle d^2 \mathbf{x}_1 d^2 \mathbf{x}_2 \\
&= \int_A \int_A e^{i \frac{k_0}{z} (\mathbf{x}_2 \cdot \mathbf{x}' - \mathbf{x}_1 \cdot \mathbf{x})} \bar{n} \delta(\mathbf{x}_1 - \mathbf{x}_2) d^2 \mathbf{x}_1 d^2 \mathbf{x}_2 \\
&= \bar{n} \int_A e^{-i \frac{k_0}{z} (\mathbf{x}_1 \cdot (\mathbf{x} - \mathbf{x}'))} d^2 \mathbf{x}_1 \\
&= \bar{n} \frac{2J(\frac{k_0 r}{z} |\mathbf{x} - \mathbf{x}'|)}{\frac{k_0 r}{z} |\mathbf{x} - \mathbf{x}'|} \\
&= \bar{n} \text{Jinc}(\frac{k_0 r}{z} |\mathbf{x} - \mathbf{x}'|)
\end{aligned} \tag{4.3}$$

where  $J$  refers to a Bessel function of the first kind and we define  $\text{Jinc}(\mathbf{x}) = \frac{2J(\mathbf{x})}{\mathbf{x}}$  and  $r$  is the radius of the source. Continuing along these lines and setting  $\mathbf{x}' = 0$  we can produce the covariance matrix of standard form as in Equation (4.1) where  $a = b = 1 + 2\bar{n}$  and  $c = d = 2\bar{n} \text{Jinc}(\frac{k_0 r}{z} |\mathbf{x}|)$ . From this, we can perform the optimisation of Equation (3.59) and obtain the explicit expressions for the mutual information, classical correlations and discord. Using the Rényi-2 entropy,

$$\mathcal{T}(\rho_{AB}) = \ln \frac{a^2}{(a^2 - c^2)}, \tag{4.4}$$

$$\mathcal{J}(\rho_{A|B}) = \ln \frac{a^2 + a}{a^2 + a - c^2}, \tag{4.5}$$

$$\mathcal{D}(\rho_{A|B}) = \ln \frac{a^2 + a^3 - ac^2}{a^2 + a^3 - ac^2 - c^2}. \tag{4.6}$$

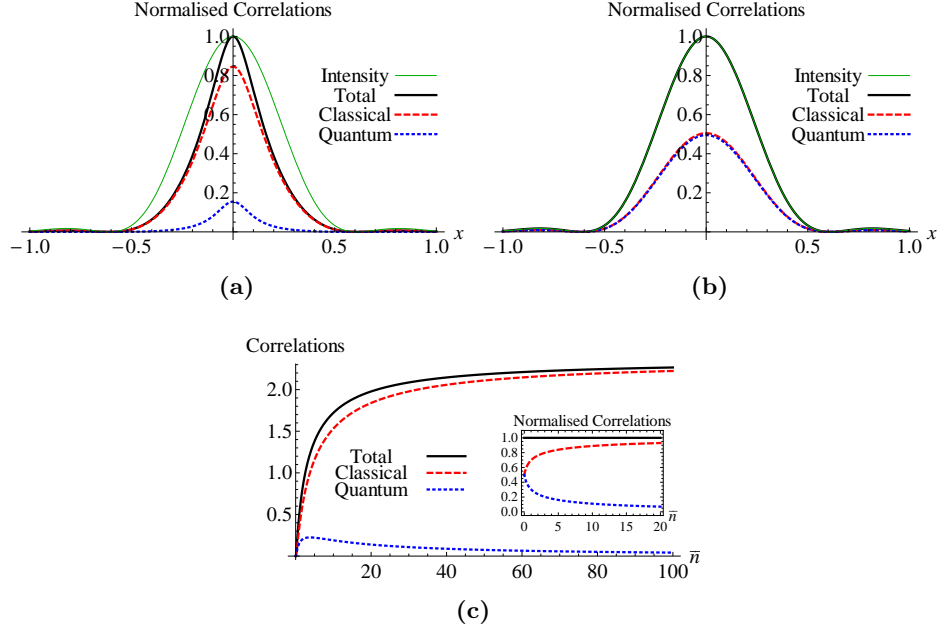
It is worth noting at this point that for the same scheme, it's a well-established result that the normalized intensity correlation function defined by

$$\mathcal{C}(\mathbf{x}, \mathbf{x}') \equiv \frac{\langle \hat{I}_1(\mathbf{x}') \hat{I}_2(\mathbf{x}) \rangle}{\langle \hat{I}_1(\mathbf{x}') \rangle \langle \hat{I}_2(\mathbf{x}) \rangle} - 1 = |\text{Jinc}(\frac{k_0 r}{z} |\mathbf{x} - \mathbf{x}'|)|^2, \tag{4.7}$$

where  $\hat{I}_i(\mathbf{x}) = \hat{a}_i^\dagger(\mathbf{x}) \hat{a}_i(\mathbf{x})$  [68]. With  $\mathbf{x}' = 0$  for one detector, we define  $\mathcal{T}(\mathbf{x})$  as the mutual information as we scan the other detector along  $\mathbf{x}$  with analogous definitions for  $\mathcal{D}(\mathbf{x})$ ,  $\mathcal{J}(\mathbf{x})$  and  $\mathcal{C}(\mathbf{x})$ . We drop the vectorial form of the transverse vector  $\mathbf{x}$  as we need only scan along one dimension due to the circular symmetry.



## 4.1.2 Results



**Figure 4.3:** Plots of the normalised total ( $\mathcal{T}$ , solid black line), classical ( $\mathcal{J}$ , dashed red line), and quantum ( $\mathcal{D}$ , dotted blue line) correlations as a function of  $x$  for (a)  $\bar{n} = 10$  and (b)  $\bar{n} = 0.01$ ; the thin green line represents the normalised intensity correlations, which practically coincides with the mutual information in (b). Panel (c) depicts the correlations as a function of  $\bar{n}$  with fixed  $x = 0$ ; the inset shows the same correlations normalised by the mutual information, to better highlight the quantum and classical contributions to the total correlations. In all the plots  $k_0 r/z = 1000$ .

To begin with, it is interesting to see how the classical and quantum correlations behave as we scan the detector for fixed values of  $\bar{n}$ . In all the following results, we have normalised the total correlations to a maximum value of 1 by plotting  $\frac{\mathcal{T}(x)}{\mathcal{T}(0)}$ ; and similarly divided quantum and classical correlations by  $\mathcal{T}(0)$ . For large  $\bar{n}$  as in Figure 4.3(a), it is apparent that the quantum correlations decay faster than the classical ones and also contribute a much smaller part of the total correlations even at their peak. However, for very small values of  $\bar{n}$  as in Figure 4.3(b) this is no longer the case: we see that both classical and quantum correlations degrade at the same rate and are approximately equal. This result stems from the fact that the purity of a Gaussian state relates only to the number of thermal photons present and for small photons counts, we obtain high

#### 4. INTENSITY INTERFEROMETRY

---

purity since the state is almost a vacuum. Recall that for pure states,  $\mathcal{D} = \mathcal{J} = \frac{1}{2}\mathcal{T}$ .

To better understand the behaviour of the correlations as we vary  $\bar{n}$ , it is useful to plot changes against  $\bar{n}$  at fixed  $x$ , see Figure 4.3(c). It is apparent that the total correlations are composed of a much greater portion of classical correlations than quantum correlations at high photon-count per mode. This gels closely with our intuition: in the regime of high brightness, the quantum component of correlations becomes very small. It is thus easy to see why in certain regimes, the effect is simple to explain classically: it appears that high photon counts wash-out quantumness, increasing the mixedness and simultaneously quashing the presence of discord.

Returning to the regime of high purity (and low photon-count), we find a notable result. The total correlations as quantified by  $\mathcal{T}$  match almost exactly the normalised intensity correlations of Equation (4.7). Notice once again that classical and quantum correlations (discord) both contribute in equal halves to such total correlations, and thus to the manifestation of the HBT effect itself in this regime. Very remarkably, for *any* form of cross-correlation of narrowband, thermal light, the correspondence between mutual information and normalised intensity correlations holds analytically in the weak-light regime (up to the third order in  $\bar{n}$ ).

To prove this, take the covariance matrix of the form (4.1) with  $a = b = 1 + 2\bar{n}$  and  $c = d = 2\bar{n}f(x)$ , where  $f(x)$  is any infinitely differentiable function. We wish to show that when  $\bar{n} \ll 1$  that  $\frac{\mathcal{T}(x)}{\mathcal{T}(0)} \approx f(x) = \left( \frac{\langle \hat{b}_1^\dagger(x) \hat{b}_2(x') \rangle}{[\langle \hat{I}(x) \rangle \langle \hat{I}(x') \rangle]^{1/2}} \right)^2 = \frac{\langle \hat{I}_1(x') \hat{I}_2(x) \rangle}{\langle \hat{I}_1(x') \rangle \langle \hat{I}_2(x) \rangle} - 1$ . The right-most equality is well known to hold for thermal light [68] and can be proven by Gaussian moment factoring (see Appendix A). It is easy to show that  $\left( \frac{\langle \hat{b}_1^\dagger(x) \hat{b}_2(x') \rangle}{[\langle \hat{I}(x) \rangle \langle \hat{I}(x') \rangle]^{1/2}} \right)^2 = f(x)$  arises naturally from the fact that  $\langle \hat{b}_1^\dagger(x) \hat{b}_2(x') \rangle = \bar{n}f(x)$  and  $\langle \hat{b}_i^\dagger(x) \hat{b}_i(x) \rangle = \bar{n}$  where  $i = 1, 2$ . In order to show that  $\frac{\mathcal{T}(x)}{\mathcal{T}(0)} \approx f(x)$ , it is useful first to recall that in standard form we can write  $\mathcal{T}(x) = \ln \frac{a^2}{a^2 - c(x)^2}$ . If we write  $g(x) = \frac{a^2}{a^2 - c(x)^2}$ , then we can say  $\frac{\mathcal{T}(x)}{\mathcal{T}(0)} = \frac{\ln g(x)}{\ln g(0)} = \log_{g(0)} g(x)$ . This then reduces the problem of showing that  $\frac{\mathcal{T}(x)}{\mathcal{T}(0)} \approx f(x)$  to showing that  $g(x) \approx g(0)^{f(x)}$ . The Taylor expansions of  $\frac{(1+2\bar{n})^2}{(1+2\bar{n})^2 - (2\bar{n}f(x))^2}$  and of  $[\frac{(1+2\bar{n})^2}{(1+2\bar{n})^2 - (2\bar{n})^2}]^{f(x)}$  match up to third order in  $\bar{n}$  (where we have imposed the condition  $f(0) = 1$  as a natural consequence of the beamsplitter transformation). This provides a proof that in low-light source conditions (i.e. small  $\bar{n}$ ), the normalised intensity correlations are almost exactly equal to the normalised Rényi-2 mutual information

in our setting. This provides an intriguing connection and a convenient short-cut for calculating and/or measuring experimentally these correlations.

All in all, it appears that there may indeed be some merit in understanding intensity interferometry from the perspective of discord-type correlations. Since the Rényi-2 entropy carries with it information about the purity of the state, we also find a simple explanation for the intuitive quantumness of the low intensity regime: the light is in a sense, *almost* entangled, although we shouldn't take this too literally. Any correlated pure state is entangled, and the states in question are very nearly pure, as can be seen by the figure 4.3. Conversely, as the intensity increases, so does the mixedness of the state (this is unsurprising, since we are considering thermal states). The discord decreases, and thereby we lose any trace of quantumness, since the state is both  $P$ -classical and effectively discord free. The suitability of classical light for describing this regime was rigorised as far back as 1950, even before the HBT result brought the whole discussion to the fore [111, 112].

Additionally, it appears that, when the intensity is low, almost all the interesting spatial information is also captured by the mutual information etc. since the spatial profile precisely matches that of the normalised intensity correlations. We shall deepen this understanding in the more intricate setting of ghost imaging.

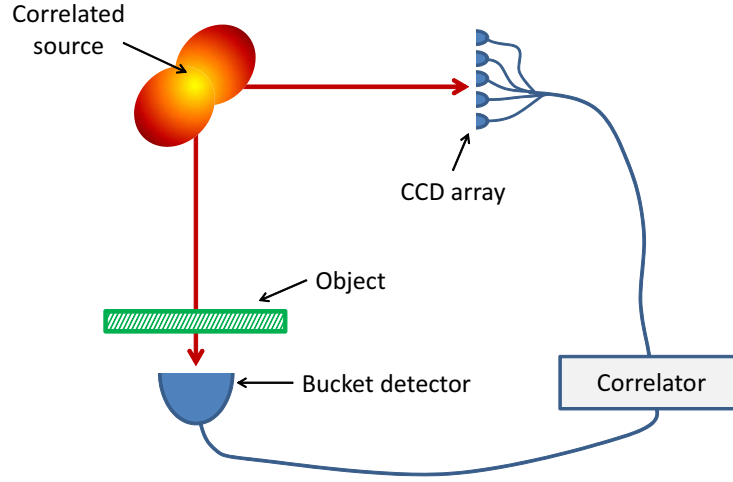
## 4.2 Ghost Imaging

Ghost imaging, is an imaging modality configured as in Figure 4.4. Behind the object to be imaged, a bucket detector is placed. This detector yields no spatial information about the object but merely informs of the net photon count or intensity of the light impinging upon it. On its own it cannot provide an image of the object. In the second branch of the scheme, however, there is a spatially resolving detector (e.g. a CCD array); in this case, the light impinging upon this detector never passes through the object so it also cannot, on its own, yield an image of the object. In order to obtain an image, it is necessary to correlate the outputs of these detectors. Additionally, we need a source of correlated light to provide the illumination to begin with; it is the exact nature of this source which has been the cause of controversy.

This controversy centred on whether or not ghost imaging could afford a classical interpretation. The boundaries of this debate continuously shifted towards an affirma-

#### 4. INTENSITY INTERFEROMETRY

---



**Figure 4.4: Diagrammatic representation of a ghost imaging scheme**

tive answer to this question; initially it was thought that the phenomenon in any form necessitated entanglement [113], but this result was elegantly disproven [114]; later, it was asserted that for the particular case of pseudothermal light it was necessary to have a quantum interpretation [115]; later still, it was suggested that this didn't apply to all pseudothermal light schemes, but only to lensless ones [106, 107], and this is where the debate lay when this work was started. In the opinion of this author, the shifting boundaries were indicative of the fact that ghost imaging is indeed well-described by semi-classical optics and as evidence to this effect increased, the debate narrowed as the proponents of the quantum interpretation had decreasing ground to stand on. A rigorous resolution of this debate was suggested by Boyd and Shapiro [26], although it was not universally accepted [116]. However, it does effectively account for the shape of the debate with a full account of suggested quantum and semi-classical interpretations and detailed, explicit physical models of each. It also contains a series of 'metalessons' which justify the work done here. Summarising the most important to us,

“

1. Semiclassical and quantum photodetection may yield identical statistics,
2. Light is intrinsically quantum mechanical, and high-sensitivity photodetection is a quantum measurement capable of revealing non-classical features in its illumi-

nation. Therefore, all optical imaging phenomena are fundamentally quantum mechanical.”

Or in other words, we *know* that semi-classical optics is entirely valid for modelling the phenomenon if the  $P$ -distribution is well-behaved (unless one intends to deny such a widely accepted and foundational result in quantum optics). However, the existence of a semi-classical interpretation does not mean that there is no quantumness in the system at all but merely that we can account for it with a hidden variable model, and it thus seems that discord, which is quantumness in a rather general sense, may somehow capture and even quantify, the unseen quantumness; our analysis of the HBT experiment provided a proof-of-principle for this. The inequivalence of bipartite quantumness as captured by the  $P$ -distribution and by discord has been shown by Ferraro and Paris to be maximal, in the sense that the sets of states which are CC-classical (as in Equation 2.24) and states which are  $P$ -classical have almost no overlap [117]. We thus expect to find new information by considering the behaviour of discord in the system.

### 4.2.1 History

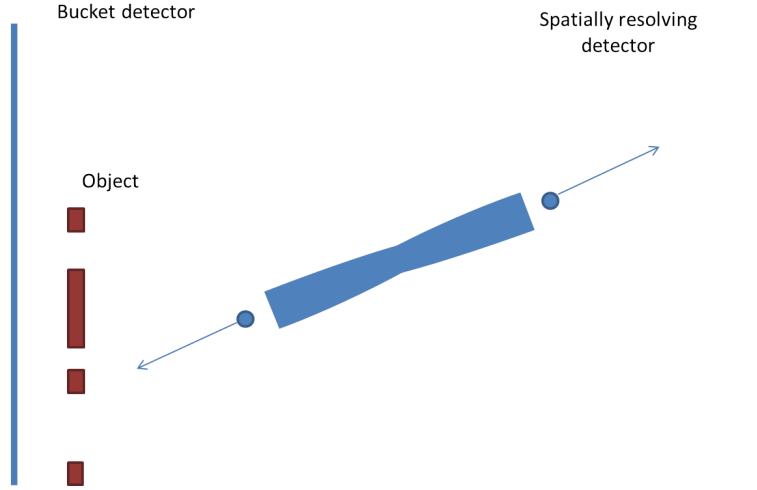
For historical completeness, as well as intuitive utility, it is worth describing the development of the ghost imaging debate chronologically. The first demonstration of ghost imaging was carried out by Pittman et al. in 1995 [100]. In their paper, they used entangled beams produced by spontaneous parametric down conversion (SPDC) to perform the imaging. Even then, they noted briefly at the end of their paper that it may be possible to adapt their technique to classical light – a proposition which sparked debate.

Abouraddy et al. [113] argued that no classical joint probability distribution between the two arms of the ghost imaging scheme could possibly produce useful information. For simplicity, we picture the light as correlated pairs of photons, one in each arm. Then Abouraddy et al.’s argument hinges on the fact that the single-photon detection probability in the spatially resolving plane  $\text{pr}_s(\mathbf{x}_s)$  must differ from the marginal probability  $\overline{\text{pr}}_s(\mathbf{x}_s) = \int \text{pr}_s(\mathbf{x}_s, \mathbf{x}_r) d^2\mathbf{x}_r$ , where the integration occurs over the entirety of the reference (bucket) plane. With a classical probability distribution, it may appear that this is clearly not possible.

#### 4. INTENSITY INTERFEROMETRY

---

However, an experimental counter-example to this argument was produced by Ben-nink et al in 2002 [114]. The error of [113] can be understood by noting that we effectively apply a filter to the detection dependent on whether or not we find a click in the bucket plane (i.e. if the photon in that plane was obstructed or not) and thus the marginal probability is not a meaningful quantity, since the choice of whether or not we count a detection is gated in this way. In other words, the bucket detector does not integrate over *all* possible outcomes because the object to be imaged adds loss to the bucket arm. A completely classical ghost imaging analogue can be conceived as in Figure 4.5.



**Figure 4.5: Completely classical ghost imaging analogue.** A double-ended cannon launches projectiles at opposing detectors. In front of a bucket detector lies an object with some gaps through which the projectile may pass. The orientation of the cannon is not directly observable, but the trajectory of both projectiles can be deduced from the position of impact on the spatially resolving detector. When the bucket detector clicks, it can be deduced that the projectile passed through one of the gaps. After a number of launches, the profile of the object can be built up in this way.

Over the following years, even more theoretical and experimental papers ironed out the differences between ghost imaging with quantum and classical light. Notably for our purposes, Gatti et al. proposed thermal-light ghost imaging which was experimentally achieved shortly afterwards [110]. This prompted a debate on the big question of whether the physical origin of thermal-light ghost imaging can be explained entirely using classical intensity correlations.

The classical interpretation of thermal ghost imaging is that it comes from ‘speckle’ correlations i.e. the correlations of the intensity profile of classical light in each arm of the scheme. The total intensity of the light impinging upon the bucket detector acts as weight function for the speckle pattern detected by the CCD, and gradually the image is formed [118]. On the other hand, the quantum interpretation posits that under certain conditions, we need to consider non-local interference effects. The indistinguishability of the photons hides the which-way path information thus creating a non-local correlation.

#### 4.2.1.1 Speckle or non-local correlations?

We follow the development of Boyd and Shapiro [26]. The field operators  $\hat{E}_s(\mathbf{x}_s, t)$  and  $\hat{E}_r(\mathbf{x}_r, t)$  are parameterised by their transverse coordinates  $\mathbf{x}_{s,r}$  and time  $t$ . We have again chosen to use the mode operators  $\hat{E}$  instead of  $\hat{a}$  for easy analogy with the corresponding classical field  $E_s(\mathbf{x}_s, t)$  (and since conventionally, the time index is often considered separately of  $\hat{a}$  in the field expansion, as in Equation (3.1)).

In experiments, the pseudothermal source is usually generated by shining a coherent beam through a rotating ground glass diffuser. The ground glass diffuser has a spatially varying phase profile, which breaks the coherence of the beam (and if bright enough, induces a speckle pattern, much like a laser pointer on a wall). The rotation ensures that the speckle pattern changes over time. Thus far we haven’t specified precisely what exactly it *means* to be pseudothermal [119], as opposed to merely thermal. A pseudothermal beam has the same temporal and spatial statistics as a thermal one, however, it is not cross-spectrally pure, which means that there may be a coupling between the temporal and spatial correlations. For our purposes, since we are only particularly interested in the spatial correlation, the two are essentially equivalent and the pseudo- qualifier an experimental technicality which we can ignore.

We also consider a transmission mask,  $T(\mathbf{x})$ , which describes the object to be imaged. The function of interest is  $\langle C(\mathbf{x}) \rangle = \langle i_r(t) i_s(\mathbf{x}, t) \rangle$ , where  $i_s(\mathbf{x}, t)$  and  $i_r(t)$  refer to the detector photocurrents for the CCD at coordinates  $\mathbf{x}$  and the entire bucket detector respectively. For semi-classical photodetection,  $i_{r,s}(t)$  are inhomogeneous Poisson processes with the rate determined by  $P_{r,s}(t) = \int_A d\mathbf{x} |E_{r,s}(\mathbf{x}, t)|^2$ , where  $A$  is the area of the bucket detector or CCD pixel in question. We have assumed units where the electron charge is 1, and also assumed perfect detectors. Clearly  $I_{r,s} = |E_{r,s}(\mathbf{x}, t)|^2$ , is

#### 4. INTENSITY INTERFEROMETRY

---

simply the classical intensity at coordinate  $\mathbf{x}$  and time  $t$ . The imaging function, for an averaging time  $\tau$  is then given by [26]

$$\begin{aligned}\langle C(\mathbf{x}_s) \rangle &= \left\langle \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} dt i_r(t) i_s(\mathbf{x}_s, t) \right\rangle \\ &= \frac{1}{\tau} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} dt \int_{A_r} d\mathbf{x}_r \langle I_r(\mathbf{x}_r, t) I_s(\mathbf{x}_s, t) \rangle |T(\mathbf{x}_r)|^2.\end{aligned}$$

As mentioned, it is well established that a time-varying ‘speckle’ pattern occurs as a result of spatial incoherence [120]. Thus both  $I_r(\mathbf{x}_r, t)$  and  $I_s(\mathbf{x}_s, t)$  possess such patterns (and identically so, since identical fields exit the beam splitter), and the term  $\langle I_r(\mathbf{x}_r, t) I_s(\mathbf{x}_s, t) \rangle$ , is the speckle correlation which forms the image.

We can identify the intensity correlation with the unnormalised form of the  $g^{(2)}$  from earlier,  $G^{(2)} = \langle E_r^*(\mathbf{x}_r, t) E_s^*(\mathbf{x}_s, t) E_r(\mathbf{x}_r, t) E_s(\mathbf{x}_s, t) \rangle$ . This is also the important component for the promised non-local interpretation.

We will discretise the source into a large number,  $M$ , of statistically independent, zero-mean modes. This allows us to approximate the fields at the detector planes by

$$\begin{aligned}E_r(\mathbf{x}_r, t) &\approx T(\mathbf{x}_r) \sum_{m=1}^M E_m(\mathbf{x}_r, t) \\ E_s(\mathbf{x}_s, t) &\approx \sum_{m=1}^M E_m(\mathbf{x}_s, t)\end{aligned}$$

where we have propagated the individual modes  $m$  from the source via  $E_m(\mathbf{x}, t) = E_m(\mathbf{x}_m, t - \frac{L}{c}) g_L(\mathbf{x}, \mathbf{x}_m) \pi r_0^2$ . Here  $g_L(\mathbf{x}, \mathbf{x}_m)$  is a propagator appropriate for distance  $L$  (so for example, in the far field, it would be the Fraunhofer propagator as with the HBT effect),  $c$  is the speed of light and  $r_0$  is the coherence length at the source, so roughly speaking, the average radius containing a mode. Recalling the large  $M$  assumption, we can ultimately make the approximation

$$G^{(2)}(\mathbf{x}_r, t; \mathbf{x}_s, t) = \left\langle \sum_{m=1}^M \sum_{m'=1}^M \left| \frac{T(\mathbf{x}_r)}{\sqrt{2}} [E_m(\mathbf{x}_r, t) E_{m'}(\mathbf{x}_s, t) + E_{m'}(\mathbf{x}_s, t) E_m(\mathbf{x}_r, t)] \right|^2 \right\rangle \quad (4.8)$$



With the assumption of low photon flux, we can interpret the contributions of each  $E_m$  term as a single photon amplitude. The above equation for points  $\mathbf{x}_r$  and  $\mathbf{x}_s$ , is formed by a sum of non-local superpositions of one term consisting of photon amplitudes from the  $m$ th coherence area and  $m'$ th coherence area of the source respectively with a second term, due to photon indistinguishability, in which  $m$  and  $m'$  are switched. Thus we indeed find that a non-local interpretation can be produced for thermal light ghost imaging.

It is worth noting at this point that we have considered a general free-space propagator  $g_L(\mathbf{x}, \mathbf{x}_m)$ . This is important because while some authors acknowledged that lensed ghost imaging, where there is a simple one-to-one correspondence between modes in the source and signal planes, could afford a classical explanation, they insisted that lensless imaging could not [116]. Since we have made no restrictions on the propagator though, we can see that this is not the case.

### 4.2.2 Correlations in ghost imaging

The existence of the semiclassical interpretation for ghost imaging is undeniable (despite the deniers), as it had to be due to the well behaved  $P$ -distribution for thermal states. However, there is still something intuitively unsatisfying with considering both low-illumination and high-illumination thermal ghost imaging as equally classical. This intuitive discomfort is what caused the debate on intensity interferometry to flare up 50 years after it was initially conceived.

It might seem easy to dismiss these intuitive misgivings since we are taught when learning quantum theory that intuition can be misleading, and rightly so. However, it is not the case that the disagreement hinges on intuition alone: as we just showed, when we assume low-illumination conditions a picture of non-local two photon interference quite easily arises.

It may well be that examining the quantum correlations in the very generalised sense of discord could help provide some succour to those who insist on the quantum picture. Despite the large number of studies addressing practical questions such as signal-to-noise ratio (SNR), image contrast and acquisition time [108, 121, 122, 123], as well as arguments based on semiclassical ideas [124, 125, 126], there has been a lack of study on the specific decomposition of classical and quantum correlations in the aforementioned sense.

## 4. INTENSITY INTERFEROMETRY

---

We have already mentioned that  $P$ -classical and discord-classical states are in a sense disjoint (specifically  $P$ -classical states are nowhere dense in the set of  $CC$ -states) and it's worth also pointing out that for bipartite Gaussian states,  $P$ -classicality implies separability. On the other hand, discord is non-zero for *any* correlated Gaussian states, which very much bolsters the case for its consideration, especially since ghost imaging relies entirely on the correlations between each arm, and discord is a correlation-like quantity, whereas the  $P$ -representation is not as intrinsically related to correlations.

Since all correlated Gaussian states contain quantum correlations in the sense of discord [94], we will consider both lensed and lensless ghost imaging, even though only the lensless variant was considered controversial at the time this work was produced. We will show that the two variants are essentially equivalent, and thus we perform most of our computations using the lensed picture, since it is simpler. We will also compare the correlations to the signal-to-noise ratio (SNR) as a useful figure of merit for the scheme. We will see that there is a high degree of similarity between the behaviour of mutual information and the SNR. It is also worth considering strictly non-classical correlations in the stronger sense of entanglement. The original ghost imaging experiment used light produced by spontaneous parametric down conversion (SPDC). This process produces none other than the two-mode squeezed states (TMSS) discussed in the introduction.

### 4.2.2.1 Lensed Ghost Imaging

We recall the schematic for a general ghost imaging scheme in Figure 4.4. In a lensed scheme, we have a simple point-to-point correspondence between modes at the detector planes and modes at the source. This is equivalent to considering a ghost imaging problem where the source lies in the far-field (at an effectively infinite distance). We will consider a discrete number of modes for simplicity.

It is necessary to clarify which variables influence the light correlations and the SNR simultaneously for our lensed picture. To begin we have the speckle-count per pixel,  $M$  of the spatially resolving detector. The second parameter of interest is the illumination  $I = M\bar{n}$ , where  $\bar{n}$  denotes the average photon count per mode—which for our lensed formulation corresponds to the photon count per speckle (i.e. we identify each mode with a speckle). The final parameter of interest is the pixel count of the image,  $R$ , which informs us of the ratio of bucket-size to pixel-size, meaning that the

bucket detector integrates over  $R \cdot M$  modes. More details on this formulation of ghost imaging can be found in [127], from which we have taken the SNR expression.

We assume that all the modes at the source have identical statistics and moreover we take the paraxial approximation. Since there is a one-to-one correspondence of modes between the planes, and we lose precise spatial information of individual modes when integrating over the pixels, the important quantity is the number of modes  $M$  and thus we suppress the position index. For the thermal-light case we can model the speckled beam as a set of independent, thermal modes  $\hat{a}_i$ . Performing the beam splitter transformation then leads to new operators  $\hat{b}_{1,i} = (\hat{a}_i + \hat{a}_{vac,i})/\sqrt{2}$  and  $\hat{b}_{2,i} = (\hat{a}_i - \hat{a}_{vac,i})/\sqrt{2}$ . The statistics for each of these beams individually are identical. The SPDC-entangled case, which we will also check for completeness, requires a very similar calculation. The difference is that our modes are generated by the usual equations  $\hat{b}_{1,i} = U\hat{a}_1(k_i) + V\hat{a}_2^\dagger(-k_i)$  and  $\hat{b}_{2,i} = U\hat{a}_2(k_i) + V\hat{a}_1^\dagger(-k_i)$ , where  $U^2 - V^2 = 1$ .

The SNR is defined as

$$\text{SNR} = \frac{|\langle S_{in} - S_{out} \rangle|}{\sqrt{\text{var}(S_{in} - S_{out})}}, \quad (4.9)$$

where our object is a binary amplitude mask; that is, we either have full transmission or full occlusion of the incoming light so  $T(\mathbf{x}) \in \{0, 1\}$ .  $S_{in}$  corresponds to being in the object profile (full transmission), and  $S_{out}$  the opposite. Here  $S$  is some general signal function, for example it may be  $g^{(2)}$  or its unnormalised form  $G^{(2)}$  as previously defined. We choose to take our signal function to be the correlation function of intensity fluctuations (the covariance) we have  $S(\mathbf{x}) = \langle (\mathbb{N}_1 - \langle \mathbb{N}_1 \rangle)(N_2(\mathbf{x}) - \langle N_2(\mathbf{x}) \rangle) \rangle$ . Here  $N_2(\mathbf{x})$  refers to the photon count of the pixel at  $\mathbf{x}$  in the spatially-resolving arm and  $\mathbb{N}_1$  denotes the net photon count on the bucket detector. To clarify, we have  $\mathbb{N}_1 = \sum_{j=1}^R N_{1,j}$ , where  $N_{1,j}$ , refers to the photon count on the  $j^{th}$  spatial resolution cell (i.e. the point on the bucket detector complementary to the pixel at  $N_2(\mathbf{x}_j)$ ). In turn we can decompose each of our  $N_1, N_2$  into the sum of individual modes (speckles) such that  $N_i = \sum_{j=1}^M n_{i,j}$  where  $i = 1, 2$ . This is the sum of photon counts for the  $M$  modes impinging on each pixel.

We then have, by inserting into Equation (4.9) and using Gaussian moment factoring

## 4. INTENSITY INTERFEROMETRY

---

yet again [127]

$$\text{SNR}_{\text{thermal}} = \frac{\bar{n}\sqrt{M}}{\sqrt{\bar{n}^2(2MR + M + 6) + 4\bar{n}(MR + 1) + 2MR + 1}}, \quad (4.10)$$

$$\text{SNR}_{\text{entangled}} = \frac{\sqrt{\bar{n}(\bar{n} + 1)M}}{\sqrt{\bar{n}^2(2MR + M + 6) + \bar{n}(2MR + M + 6) + 1}}, \quad (4.11)$$

for thermal-light and SPDC-entangled light sources, respectively.

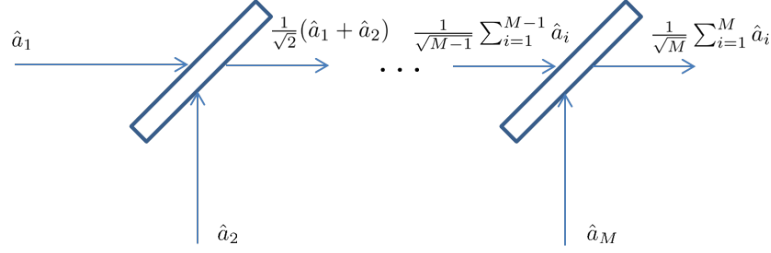
We identify two regimes to explore. The first is the mode-to-mode (speckle-to-speckle) correlations. These depend only upon  $\bar{n}$  and thus, given the above equations for SNR, clearly do not capture much of the actual physics of the system. However, we will also consider a two-mode approximation of the entire system in order calculate the bucket-to-pixel correlations.

### 4.2.2.2 Coarse-grained correlations

In order to have useful spatial resolution, it is desirable to have a large number of modes. This renders the problem of calculating discord and classical correlations intractable. We thus wish to define effective operators over entire areas corresponding to pixels of the detectors. The key to this is to produce averaged operators which under the expectation value behave like the usual single-mode operators.

To wit, starting with a pixel which collects  $M$  modes and noting that  $N_i = \sum_{j=1}^M n_{i,j} = \sum_{j=1}^M \hat{b}_{i,j}^\dagger \hat{b}_{i,j}$  (for  $i = 1, 2$ ), we wish to find mode operators  $\hat{c}_i$  for each pixel, such that  $\langle \hat{c}_i^\dagger \hat{c}_i \rangle = \langle \sum_{j=1}^M \hat{b}_{i,j}^\dagger \hat{b}_{i,j} \rangle$ . In order to do this, we simply take a linear combination of the modes. Doing so, we can say  $\hat{c}_i = \sum_{j=1}^M \hat{b}_{i,j}$ . Note that we do not consider the effects of the transmission mask on light propagation in this particular calculation.

These operators then behave as we desire, but for one very crucial part: the commutation relations. This is easily remedied though. If we define our effective operators in the following manner:  $\hat{d} = \frac{1}{\sqrt{M}} \sum_{i=1}^M \hat{c}_i$ , then we find  $[\hat{d}, \hat{d}^\dagger] = 1$ . Operationally, this definition corresponds to taking an arrangement of beam-splitters which combine all the modes with equal weight, so it indeed corresponds to a physical averaging of modes (Figure 4.6), though an experimentally infeasible one due to their spatial distribution. Performing similar steps for pixels collecting  $R \cdot M$  modes, and then transforming our



**Figure 4.6:** An ‘ $M$ -splitter’ consisting of  $M - 1$  beam splitters where the  $k$ th beam splitter from the left has a transmissivity of  $\frac{1}{\sqrt{M-k+1}}$ , thus producing an equal combination of all modes.

covariance matrices from the mode operator to the quadrature basis, enables us to obtain the effective ‘coarse-grained’ covariance matrices given by

$$\begin{pmatrix} 1 + 2\bar{n} & 0 & \frac{2\bar{n}}{\sqrt{R}} & 0 \\ 0 & 1 + 2\bar{n} & 0 & \frac{2\bar{n}}{\sqrt{R}} \\ \frac{2\bar{n}}{\sqrt{R}} & 0 & 1 + 2\bar{n} & 0 \\ 0 & \frac{2\bar{n}}{\sqrt{R}} & 0 & 1 + 2\bar{n} \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 + 2\bar{n} & 0 & \frac{2\sqrt{\bar{n}+\bar{n}^2}}{\sqrt{R}} & 0 \\ 0 & 1 + 2\bar{n} & 0 & \frac{2\sqrt{\bar{n}+\bar{n}^2}}{\sqrt{R}} \\ \frac{2\sqrt{\bar{n}+\bar{n}^2}}{\sqrt{R}} & 0 & 1 + 2\bar{n} & 0 \\ 0 & \frac{2\sqrt{\bar{n}+\bar{n}^2}}{\sqrt{R}} & 0 & 1 + 2\bar{n} \end{pmatrix}, \quad (4.12)$$

for thermal-light and SPDC-entangled light cases, respectively.

We will also find that, primarily for aesthetic purposes, it is convenient to normalise the correlations by multiplying them by  $\sqrt{\frac{R}{2}}$ , which will ensure the mutual information and SNR vary almost identically in all three parameters of interest. We can provide a partial justification of this by noticing that the off-diagonal blocks of the covariance matrices, which encode the intermodal correlations, appear scaled by a factor  $\propto 1/\sqrt{R}$  as a consequence of the averaging applied to preserve the commutation relations. This makes e.g. the covariance  $\langle \hat{a}_{A_1}^\dagger \hat{a}_{A_2} \rangle$  sensitive to changes in one of the areas, which is not particularly desirable, as an increase in, say,  $A_2$  for a fixed  $A_1 \leq A_2$  implies no loss of any physical correlation. Therefore, in the effective two-mode description, we can *a posteriori* renormalise classical, quantum, and total correlations<sup>1</sup>, evaluated on the matrices of Equation (4.12), multiplying them by a factor proportional to  $\sqrt{R}$ . The specific choice  $\sqrt{R/2}$  is dictated by mere convenience, as we will see it makes the slope in

<sup>1</sup>In this section we actually use the von Neumann entropy, as the work was completed before the development of the framework of Rényi entropy correlations. The results are qualitatively the same, either way.

#### 4. INTENSITY INTERFEROMETRY

---

the SNR versus total correlations relation converge exactly to 1 [see Figure 4.10(c) and Equation (4.13)]. It's also true that the quasi-linear interdependence existing between the two quantities for thermal light, as well as all the upcoming results illustrated in Figures 4.10 and 4.11, are obviously not qualitatively affected (albeit for quantitative rescaling) by any specific normalisation procedure implemented on the correlations.

At this point we will show that we can coarse-grain the lenseless case in precisely the same manner as the lensed one. While lensed ghost imaging considers a simple one-to-one mapping from the operators in the source plane to the ones at the detection planes, we can generalise our calculations to other sorts of propagation. For the purposes of lensless ghost imaging, it is conventional to consider free-space propagation in the paraxial approximation. In this case, it is possible to characterise the field at transverse point  $\mathbf{x}_j$  in plane  $z_j$  by  $\hat{E}(\mathbf{x}_j, z_j; k) \propto \sum_k g(\mathbf{x}_j, z_j; k) \hat{a}(k)$  for  $j = 1, 2$ , where  $\hat{a}(k)$  are our source-plane operators and  $g(\mathbf{x}_j, z_j; k)$  is the Green's function describing the propagation of the field to the point with transverse coordinates  $\mathbf{x}_j$  on the detection plane [106, 128]. The form of  $g(\mathbf{x}_j, z_j; k)$  in the quasimonochromatic case is [128]:

$$g(\mathbf{x}_j, z_j; k) = \frac{-ik_0 e^{ik_0 z_j}}{2\pi z_j} \int d\mathbf{x}_s e^{i \frac{k_0}{2z_j} |\mathbf{x}_j - \mathbf{x}_s|^2} e^{-ik \cdot \mathbf{x}_s}.$$

We can then calculate the auto-correlations and cross-correlations and from this our effective covariance matrix. To begin, we note that in the quasimonochromatic, paraxial approximation,  $[\hat{E}_i(\mathbf{x}_i, z_j), \hat{E}_j^\dagger(\mathbf{x}_j, z_j)] = \delta(\mathbf{x}_i - \mathbf{x}_j) \delta_{ij}$ , where  $\mathbf{x}$  are the coordinates in the transverse plane [121, 129]. As such, these behave under the expectation value as though they were the usual single mode operators  $\hat{a}$ . Henceforth, we will assume that we are always considering correlations on planes at equal distances from the source as is conventional for ghost imaging setups. For individual modes or correlated pairs:

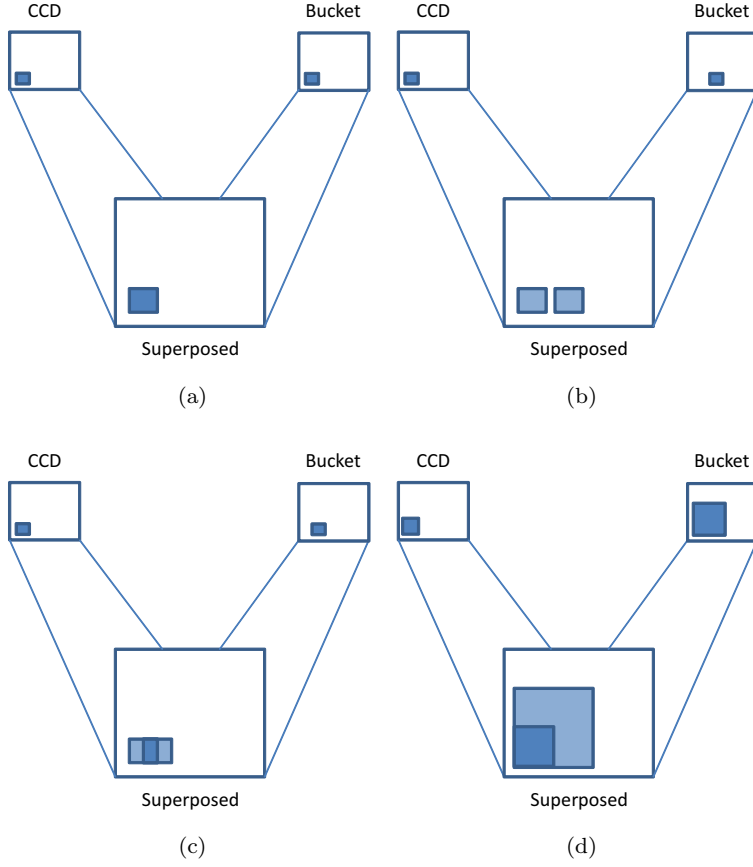
$$\begin{aligned}
 \langle \hat{E}^\dagger(\mathbf{x}_i, z) \hat{E}(\mathbf{x}_j, z) \rangle &\propto \langle \sum_{kk'} g^*(\mathbf{x}_i, z; k) g(\mathbf{x}_j, z; k') \hat{a}^\dagger(k) \hat{a}(k') \rangle \\
 &= \sum_{kk'} \langle g^*(\mathbf{x}_i, z; k) g(\mathbf{x}_j, z; k') \rangle \langle \hat{a}^\dagger(k) \hat{a}(k') \rangle \\
 &= \sum_k \left\langle \left| \frac{k_0 e^{ik_0 z}}{2\pi z} \right|^2 \int d\mathbf{x}_s e^{-i\frac{k_0}{2z}|\mathbf{x}_i - \mathbf{x}_s|^2} e^{ik \cdot \mathbf{x}_s} \int d\mathbf{x}'_s e^{i\frac{k_0}{2z}|\mathbf{x}_j - \mathbf{x}'_s|^2} e^{-ik \cdot \mathbf{x}'_s} \right\rangle \\
 &\quad \times \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle \\
 &= \left( \frac{k_0}{2\pi z} \right)^2 \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle \sum_k \left\langle \int d\mathbf{x}_s d\mathbf{x}'_s e^{i\frac{k_0}{2z}|\mathbf{x}_j - \mathbf{x}'_s|^2 - |\mathbf{x}_i - \mathbf{x}_s|^2} e^{ik \cdot (\mathbf{x}_s - \mathbf{x}'_s)} \right\rangle \\
 &\propto \left( \frac{k_0}{2\pi z} \right)^2 \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle \left\langle \int d\mathbf{x}_s e^{i\frac{k_0}{2z}|\mathbf{x}_j - \mathbf{x}_s|^2 - |\mathbf{x}_i - \mathbf{x}_s|^2} \right\rangle \\
 &= \left( \frac{k_0}{2\pi z} \right)^2 \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle e^{i\frac{k_0}{2z}(\mathbf{x}_j^2 - \mathbf{x}_i^2)} \int d\mathbf{x}_s e^{i\frac{k_0}{z}\mathbf{x}_s \cdot (\mathbf{x}_j - \mathbf{x}_i)} \\
 &\approx \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle \delta(\mathbf{x}_j - \mathbf{x}_i).
 \end{aligned}$$

For the calculation, we have assumed a large disk-like source and a large number of modes. These are standard assumptions made in the derivation of the non-local biphoton model of ghost imaging [106, 107]. Our calculation indicates that  $\langle \hat{E}^\dagger(\mathbf{x}_i, z) \hat{E}(\mathbf{x}_j, z) \rangle \propto \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle$  when  $\mathbf{x}_i = \mathbf{x}_j$ , and is 0 otherwise. The correlation at any point on a single plane or paired points on CCD and bucket planes is proportional to the expectation value for a given mode in the momentum-basis. This behaviour reflects that of lens-based ghost imaging. In fact, it shows that we have a proportionality between photon counts in the source-plane and detection planes, which is not an entirely surprising result.

We can also go further and define effective operators in a similar manner to how we did for lens-based imaging by writing  $\hat{E}_p = \frac{1}{\sqrt{A_p}} \int_{A_p} \hat{E}(\mathbf{x}, z) d\mathbf{x}$  where we are integrating over a pixel with area  $A_p$ . We can test that the correct commutation relation holds for this operator under the expectation value using similar assumptions as above. It is then also easy to show that  $\langle \hat{E}_p^\dagger \hat{E}_p \rangle \propto \langle \hat{a}^\dagger(k) \hat{a}(k) \rangle$ . Again, the behaviour reflects that of lensed ghost imaging. This leaves an integration over the bucket detector to fully characterise our scheme. The only difference to the above calculation is that we are integrating over a larger area,  $A_b$ . It is evident that the auto-correlations will be the

#### 4. INTENSITY INTERFEROMETRY

---



**Figure 4.7: Conceptual diagrams for coarse-grained correlations**, indicating the areas we are considering on the CCD and bucket planes and a superposition of both. For part (a) we are considering exactly matched areas on either plane, for part (b) we are considering entirely disjoint areas. Part (c) has two areas of equal magnitude, half overlapping. The correlations between these are  $\frac{1}{2}$  those for an individual speckle within the overlapping area. Part (d) considers one area  $\frac{1}{4}$  the size of the other, completely enclosed by it on the superposed picture. This case notably, is most like the comparison of pixel and bucket detector. The effective modes also have  $\frac{1}{2}$  the cross-correlations between them as for any individual speckles within the overlap area.

same. For the cross-correlations we find  $\langle \hat{E}_b^\dagger \hat{E}_p \rangle \propto \langle \frac{\hat{a}^\dagger(k) \hat{a}(k)}{\sqrt{\frac{A_b}{A_p}}} \rangle$ . Essentially, we scale the cross-correlations by the inverse square root of the ratio of the bucket detector area to the pixel area. This yields an identical scaling with changes in our three parameters  $I, R, M$  as for the lens-based imaging.

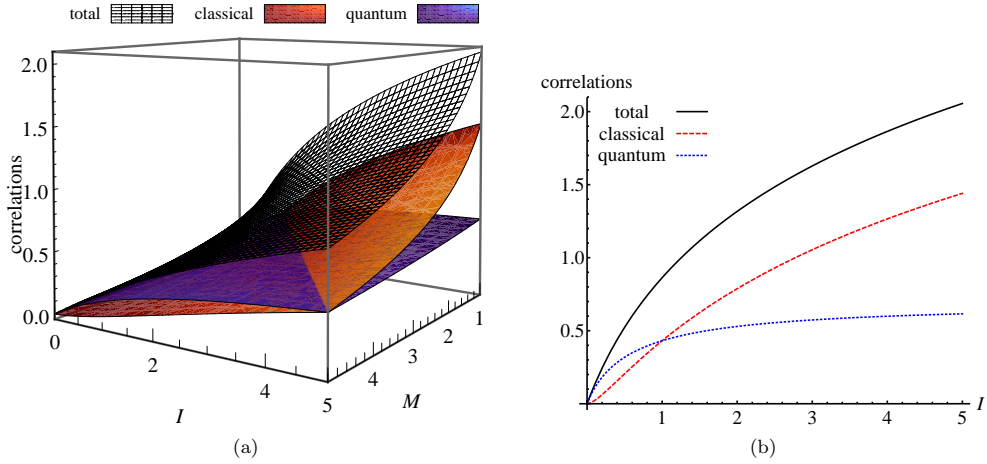
To better understand this coarse graining – and to see that it has some sensible



properties, we illustrate how the auto- and cross-correlations using the effective operators behave under several circumstances in Figure 4.7. Ultimately, we find that the autocorrelations are always the same as those for a single mode, while the *unnormalised*, coarse-grained cross-correlations scale as  $\frac{A_{\text{overlap}}}{\sqrt{A_1 A_2}}$ . Panel (d) in particular is of relevance for ghost imaging; in this case, we find that the cross-correlations scale as  $\sqrt{\frac{A_{\text{overlap}}}{A_2}}$ , or equivalently  $\sqrt{\frac{A_{\text{pixel}}}{A_{\text{bucket}}}} = \frac{1}{\sqrt{R}}$ .

#### 4.2.2.3 Results

We will temporarily relieve ourselves of the coarse-grained picture, and start by looking at the mode-to-mode correlations. As with the HBT effect, the results hint that even with non-entangled thermal light, the role of quantum correlations cannot be ignored. In Figure 4.8, we clearly see that for low illumination  $I$  – or alternatively, high speckle-count per pixel  $M$  – quantum correlations can actually *exceed* classical ones. This stems from the fact that for individual pairs, the correlations only depend on the expected photon count  $\bar{n} = \frac{I}{M}$ . When  $\bar{n} < 1$  the speckle-speckle quantum correlations dominate over the classical ones,  $\mathcal{D} > \mathcal{J}$ . Entanglement is never present in the considered light source, yet this reveals a definite non-classical nature of such light, manifested in the correlations between individual pairs of speckle modes in the low illumination regime.

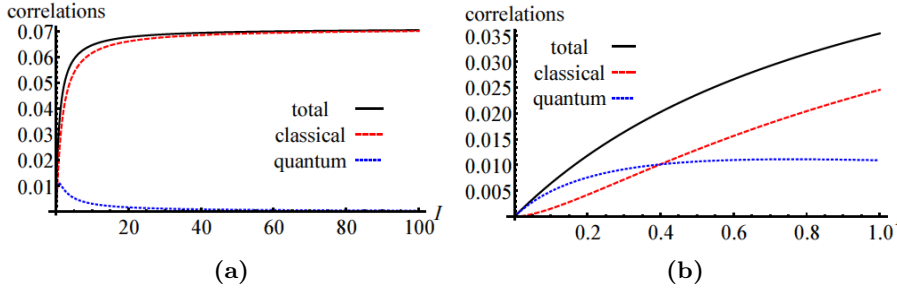


**Figure 4.8: Speckle-speckle correlations in thermal-light lensed ghost imaging,** plotted as a function of the illumination  $I$  and the speckle-count  $M$ . In (b) detail at  $M=1$  is shown. Notice the intersection between classical and quantum correlations at  $M = I$ . Classical ones dominate for  $I > M$ , while quantum ones are relevant for  $I < M$ .

#### 4. INTENSITY INTERFEROMETRY

---

The results we find for the coarse-grained correlations are even more striking. In the following, we refer to the normalised (multiplied by  $\sqrt{\frac{R}{2}}$ ) total, quantum, and classical coarse-grained correlations by  $\tilde{\mathcal{T}}$ ,  $\tilde{\mathcal{D}}$ , and  $\tilde{\mathcal{J}}$ , respectively. Setting  $R = 100$ , we find the correlations scale as in Figure 4.9.

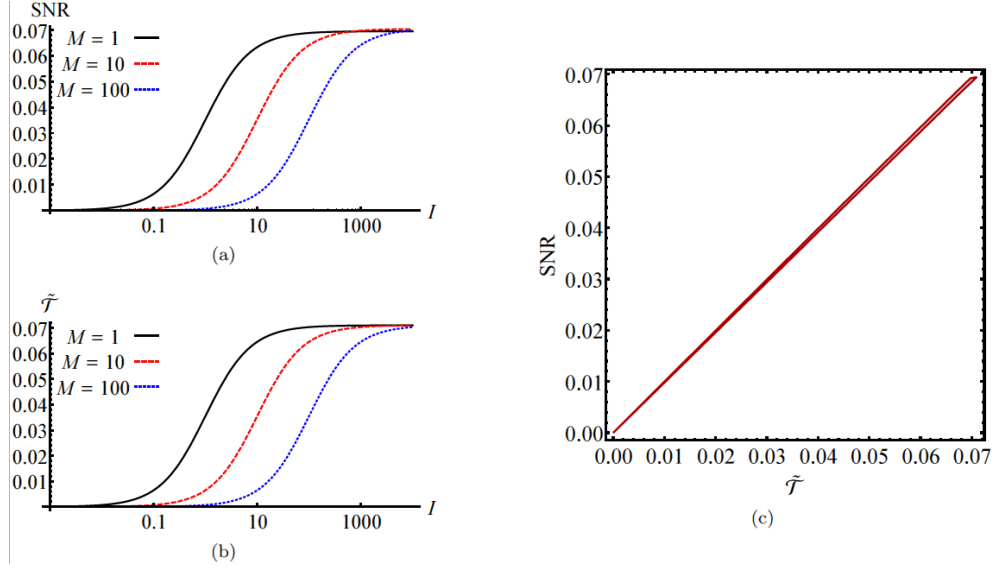


**Figure 4.9: The decomposition of coarse-grained correlations** at  $M = 1$  and  $R = 100$ . Looking at (a) we can see that the total correlations are bounded, and at high illuminations they are mostly classical. Zooming in to the low-illumination regime near the origin, though, we see in (b) that there is an interval where quantum correlations still exceed classical ones.

This is quite promising. As we hoped, there does seem to be a significant distinction between the low-illumination and high-illumination regimes. In fact, it seems fair to say that the low-illumination regime is mostly quantum, whereas the high illumination one mostly classical.

It is useful to compare the behaviour of the correlations with the SNR, which quantifies the quality of the imaging. In this way we can attempt to identify which aspects of the correlations in the source beams can, in a loose sense, be regarded as resources for the protocol.

To begin, we first plot the *total* normalised coarse-grained correlations  $\tilde{\mathcal{T}}$  and the SNR on separate graphs in Figure 4.10. Despite the very different (physical and mathematical) nature of the two quantities under scrutiny, it is immediately noticeable that they have a very similar form, a fact which is even more evident when we observe a parametric plot of the SNR versus  $\tilde{\mathcal{T}}$  in Figure 4.10(c). By varying  $I$ ,  $M$ , and  $R$  in their region of interest (in particular keeping  $R \gg 1$ , which means nontrivial imaging), we find that the SNR always exhibits a quasi-linear dependence on the total correlations with slope  $\approx 1$ . The relation is not exactly linear, yet the discrepancy between SNR and normalised total correlations stays smaller than one percent in the relevant parameter



**Figure 4.10: Panels [(a)-(b)]: Log-linear plots of (a) Signal-to-noise ratio (SNR) and (b) total correlations  $\tilde{T}$  for increasing illumination  $I$ . In panel (c) we display the tiny region filled by SNR versus  $\tilde{T}$  with varying  $M \in [1, 1000]$  and  $I \in [0, 1000]$ ; the quasi-linear region is essentially invariant upon variations of  $R$  as well.**

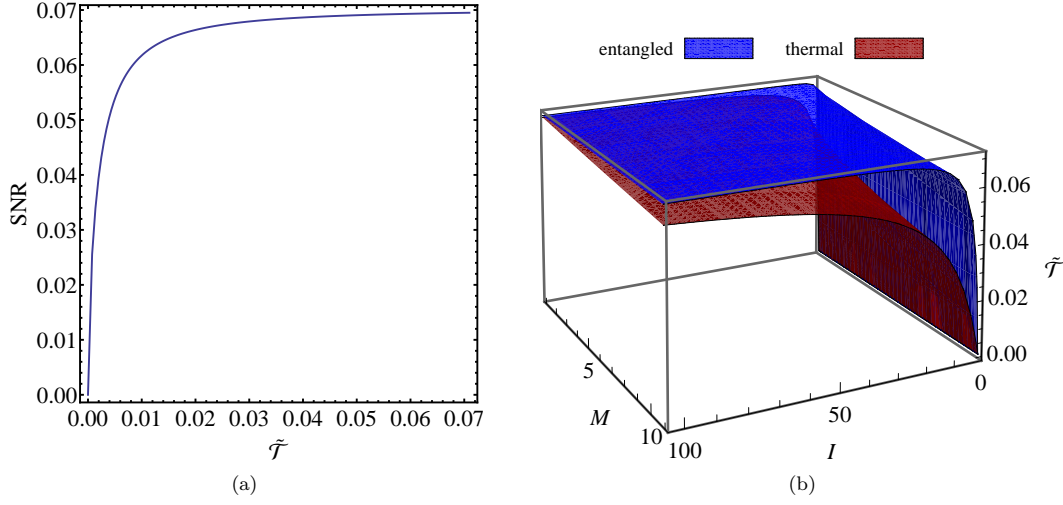
regime. This indicates that for *any* of the parameters that affect the light correlations, the SNR is affected in exactly the same way. The small difference may reflect the fact that there are properties of the detector and object which affect the SNR, but do not influence the correlations at the source. Rigorously, recalling  $\bar{n} = I/M$ , and using the formulas provided previously, we have

$$\frac{\text{SNR}}{\tilde{T}} \xrightarrow{\bar{n} \rightarrow \infty} \frac{\sqrt{1/(2R+1)}}{\sqrt{R/2} \ln[R/(R-1)]} \xrightarrow{R \rightarrow \infty} 1. \quad (4.13)$$

Our analysis thus reveals that the joint contribution of quantum *and* classical correlations can actually be used as a reasonable predictor for the performance, as measured by the SNR, of ghost imaging with ‘classical’, thermal-light sources.

We have comprehensively unveiled the presence and role of genuinely quantum correlations in thermal-light, ‘classical’-like ghost imaging. We can also extend our study to the case of ghost imaging with entangled light sources produced by SPDC. In this case, it has already been pointed out in [122] that the entanglement between the individual modes acts as an extra resource and can lead to an increase in the SNR compared to the thermal-light case. In our formalism, we find that for entangled

#### 4. INTENSITY INTERFEROMETRY



**Figure 4.11: (a) Coarse-grained analysis for ghost imaging using entangled light** produced by SPDC (with pixel count  $R = 100$ ): Plot of SNR versus total correlations for  $M = 1$  and  $I \in [0, 1000]$ ; compare with the corresponding thermal-light case, Figure 4.10(c). Panel (b) depicts a comparison between total coarse-grained correlations  $\tilde{T}$  for the cases of thermal and entangled light sources (with  $R = 100$ ) as a function of the illumination  $I$ ; it is shown that they share a common limit in the regime of high illumination [see Equation (4.13)], although in this setup SPDC-entangled light is always more correlated than thermal-light. All the correlations are normalised by  $\sqrt{R/2}$ .

light the SNR grows much faster than linearly as a function of the coarse-grained total correlations  $\tilde{T}$ , as shown in Figure 4.11(a). However, in the limit of very high illumination, the coarse-grained total correlations for both thermal-light and entangled cases converge to the same asymptotic maximum value, which is an intuitively expected result [121]. Similarly, the SNR for both cases converges to the same upper limit in the regime of high illumination, given by  $\text{SNR}_{\tilde{n} \rightarrow \infty} = \sqrt{M/[6 + M(2R + 1)]}$  [127]. Specifically, we find that Equation (4.13) holds exactly for entangled as well as thermal light. When we plot the (normalised, coarse-grained) total correlations for thermal-light versus SPDC-entangled sources on the same graph, as in Figure 4.11(b), it is evident though that in the case of entangled light the correlations are always higher for given finite values of the parameters  $I$ ,  $M$  and  $R$ .

#### 4.2.2.4 Discussion

We have analysed the nature of correlations in Gaussian light sources used for ghost imaging from a quantum informational perspective, combining a microscopic with an effective coarse-grained description. We have found that even so-called ‘classical’ thermal light contains nonzero genuinely quantum correlations – as measured by the quantum discord [54, 55] – whose contribution to the performance of ghost imaging schemes has been assessed. Since the entire scheme is dependent upon the correlations between the arms, the quasi-linearity between the total correlations and SNR strongly suggests that the total correlations may be acting as the essential operative resource. It is important to notice that this correspondence could vary depending on the choice of signal function, or if a more complex object is chosen; the model for calculating correlations is a coarse one and thus we should be wary of putting too much weight on the direct correspondence, striking though it may be.

The dominant strength of quantum correlations in low-illumination regimes has an immediate physical explanation. In the limit of low illumination, there are few photons per mode and this is a regime where the quantum behaviour of light becomes very apparent. This is consistent with the non-local picture of Boyd and Shapiro which we discussed at the start and requires explicitly the assumption of low-flux, [26] and moreover in the original paper expounding the non-local picture [107], the adopted model of thermal light implicitly assumes a photon-counting regime, which is an equivalent criterion to low photon-flux. As elucidated by the results of this paper, it is likely no mere coincidence that such an assumption need be made, but rather a consequence of the fact that the quantum component of correlations available for detection by our scheme vanishes in the limit of high illumination.

Unfortunately, the quantum discord has no simple rigorous interpretation in terms of non-local effects [58]. Although the discord provides a strong intuitive aid for understanding ghost imaging in that there is some proportionately greater level of disturbance upon measuring in the low-light regime, it does not have a simple operational interpretation we can apply here. It is also important not to overstate the ‘non-local’ picture, since the usual information theoretic understanding of non-locality involves violating a Bell inequality, which is absolutely not possible here, since this is an even stronger requirement than entanglement existing in the state. After all, the positivity

## 4. INTENSITY INTERFEROMETRY

---

of the  $P$ -representation means that a local hidden variable model is always available for describing thermal light ghost imaging.

Nevertheless, the coarse-grained formalism put forward here seems to indicate how the quantum nature of the light source becomes quenched as we diverge from the photon-counting regime and enter the classical limit of intensity correlations. For these high illuminations, the quantum correlations available for detection by our scheme tend to zero, and the physical model of the scheme does not require a quantum description of the light to be accurate. The small-scale features of the speckle-speckle correlations cannot be detected by the way the light is averaged over the bucket detector. Even though there is no explicit decoherence, we gain a clear look into a quantum system appearing to transition into a classical one due to the nature of the measurement process.

We have briefly extended our analysis to the less controversial case of SPDC-entangled light sources. It has been observed in previous work [121] that for high brightness the results of the ghost image formed with ‘classical’-like thermal light are “excellent approximations for the quantum [entangled] case” and we find indeed that in this limit, the (coarse-grained) total correlations in the sources too approximate each other closely.

Our calculations reveal furthermore that lensed and lensless ghost imaging behave in a similar manner to each other. In the debate, which is well characterised by a comment from Shih [116] with a rebuttal from Shapiro and Boyd [130], we find our results further bolster the latter’s arguments that near-field lensless imaging can afford a classical interpretation. In high illuminations the quantum correlations are suppressed. Ultimately, this shows that the geometry of the system has a limited role in determining its classicality or quantumness, whereas the illumination conditions are a clearly more significant factor.

### 4.3 Quantum Illumination

The work in the previous section was published in [1]. After its completion, the opportunity for an experimental analysis of a different intensity interferometry scheme arose, and this section will detail that work. The experiment, and some of the accompanying

theory, was carried out by the coauthors listed in [3]. The analysis in terms of the coarse-grained picture as defined above is all my own work.

Strictly speaking, the original formulation of quantum illumination did not refer to an intensity interferometric setup. Lloyd [103] was the first to conceive of a discrete-variable quantum illumination scheme, which acts effectively as a quantum radar enhancing our ability to detect an object by using entanglement. The more rigorous formulation of quantum illumination was provided by Tan et al. and used Gaussian states [104].

The rigorous formulation of quantum illumination invokes the quantum Chernoff bound. The classical Chernoff bound, an important result for hypothesis testing, is used to determine the asymptotically minimal probability of error when discriminating between two probability distributions given a series of outputs drawn from these distributions. The quantum Chernoff bound does the same for multiple copies of density matrices [131], and thus in a radar-type scheme, if we wish to determine whether an object is present or absent, it is clearly quite applicable.

Using this bound, Tan et al showed that in the case of an object with low-reflectivity sitting in a noisy bath, directing one part of a TMSS towards the object and retaining the other part – the reference beam – for joint measurement with the (potentially) reflected light, enables a lower probability of error than for any unentangled inputs. In fact, they showed that the upper bound for the probability of error with entangled light was better than the lower bound on the probability of error for unentangled light. This is feasibly useful if one desires to detect an object with as few photons possible. This could be useful for purposes of stealth, or if the object to be detected is vulnerable to damage from bright light.

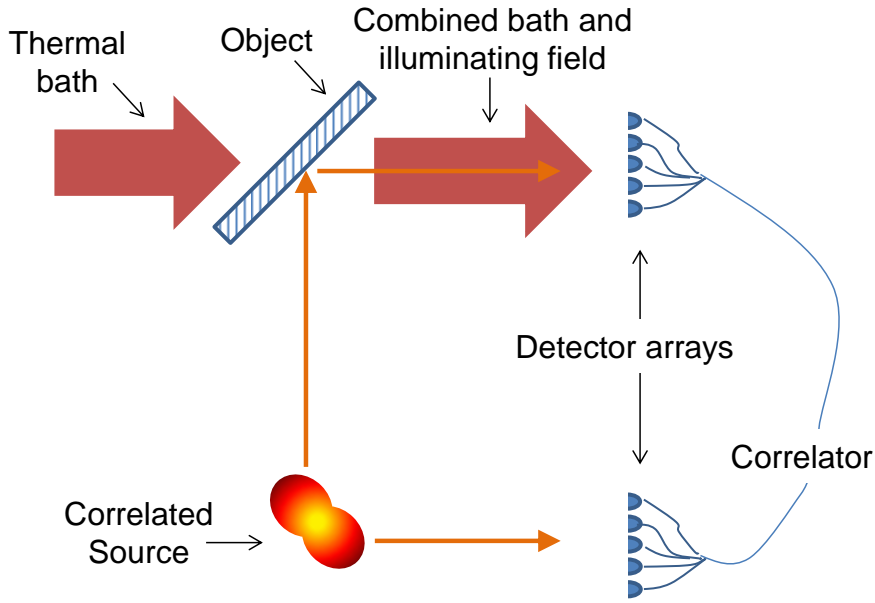
What makes quantum illumination quite remarkable, is that it is designed to operate with a large amount of noise in mind. Usually noise is considered to be the enemy of entanglement, and indeed, due to this noise the entanglement is broken between the reference beam and the reflected light, however, we still find an enhancement when using a two-mode squeezed state input, which beats any possible unentangled source. It seems then that it isn't really the entanglement that produces the enhancement but perhaps some other property of the light which is not broken by the noise.

The term 'quantum illumination' has also been used in a looser sense than that of Tan et al., and in [9] was used to refer to an intensity interferometry setup. The

#### 4. INTENSITY INTERFEROMETRY

---

analysis of [9] is not performed with regards to the Chernoff bound as in the rigorous formulation, but simply considers that *given a certain measurement scheme* – in this case an intensity interferometric one – an entangled input performs better than an unentangled one. This is still useful, since the scheme is much more easily implemented than known schemes which violate the Chernoff bound for classical light, and moreover there are no known schemes whatsoever which actually saturate the Chernoff bound for entangled beams anyway.



**Figure 4.12: A schematic diagram of a quantum illumination scheme.** Here the object under consideration is a beam splitter. A portion of the light produced by the correlated source will be reflected by the beam splitter and the challenge is to detect this beneath the dominant thermal noise, thereby discriminating the presence of the object.

This intensity interferometric quantum illumination scheme which we will apply our analysis to is presented in Figure 4.12. In the more general quantum illumination, we do not need to consider CCD detector arrays, but may have any sort of measurement whatsoever.

Aside from the measurement scheme, there are other non-trivial differences between the quantum illumination scheme we consider and that of Tan et al. Tan’s scheme involves the detection of a single mode which interacts with the noisy background, whereas our scheme is intrinsically multi-mode and the model (experimentally tested



as accurate) assumes that the signal and noise modes do not combine with each other on the beam splitter (but rather, both combine with vacuum modes). The multi-mode setup, as well as the large number of pixels on the CCD enable a detection test to be performed in one shot, whereas the Chernoff bound techniques require a large number of tests. As such, the scheme we consider is really quite a pragmatic one, even if it does not have the same theoretical appeal as the original.

The multi-mode structure also renders the intensity interferometric quantum illumination variant well suited for applying the techniques we’ve developed thus far. While it is not possible to use the original quantum illumination formalism for this, by using quantumness criteria related to the Glauber Sudarshan  $P$ -representation, it is possible to show that a quantum input in this scheme is largely more powerful than any classical one with the same local statistical properties, in terms of photon number detection and correlation measurements. As in the case of [103, 104], in the scheme of [9], entanglement is completely destroyed before the detection stage.

#### 4.3.1 Setup and analysis

In the experimental setup implemented in Ref. [9] the object to detect is a 50:50 beam splitter (BS) embedded in a “bath” of thermal modes. The light source used to probe the presence of the object consists of multiple identical and independent pairs of either two mode squeezed state twin beams (TWB) or correlated, unentangled thermal beams (THB). Charge-coupled device (CCD) arrays are placed in each of the signal and reference planes, as outlined in the scheme of Figure 4.12. As with ghost imaging, each pixel in the signal plane collects  $M$  TWB (or THB) modes resulting in a net photon count per pixel  $N_I$  (excluding the bath), which are correlated with  $M$  corresponding modes intercepting another pixel in the reference plane counting  $N_r$  photons (see also [132, 133]). We denote by  $N_\beta$  and  $M_\beta$  the photon count and the number of modes per pixel of the bath. The losses in each arm are taken into account by the detection efficiencies  $\eta$  for the illuminating light (where this quantity does not include the non-unit reflectivity of the object in the reference plane) and  $\eta_\beta$  for the bath.

The covariance of photon counts per pixel on the signal plane  $N_s = N_I + N_\beta$  and on the reference plane  $N_r$  is evaluated averaging over the set of pixel pairs  $N_{pix}$  in one shot of the CCD ( $N_{pix} = 80$  in the experiment). The covariance is expected to vanish when the object is absent, so the detection of the target is declared when the

## 4. INTENSITY INTERFEROMETRY

---

covariance is larger than a certain threshold value. The SNR (normalized by  $N_{pix}^{1/2}$ ) for the measurement, given by the difference of covariances in the cases when the object is present ( $\Delta^{\text{in}}$ ) or absent ( $\Delta^{\text{out}}$ ), is [9]

$$\text{SNR} = \frac{|\langle \Delta^{\text{in}} - \Delta^{\text{out}} \rangle|}{\sqrt{\delta^2 \Delta^{\text{in}} + \delta^2 \Delta^{\text{out}}}} \quad (4.14)$$

for  $\Delta = N_s N_r - \langle N_s \rangle \langle N_r \rangle$ , where  $\delta^2 X = \langle X^2 \rangle - \langle X \rangle^2$ . Remarkably, the TWB entangled input, representing the maximal allowable cross-correlation between two modes, harbours a hefty improvement over the case of maximally correlated modes within the bounds of separability i.e. with a proper  $P$ -representation, as experimentally demonstrated in [9, 104].

The complete experimental setup we are investigating is extremely complicated due to the presence of many modes and high levels of noise. As with ghost imaging, this makes it infeasible to pursue a direct analytical approach to the problem. We can overcome this difficulty by invoking our effective-mode description. We now have to consider the bath and thus the modes in the signal plane will be given by

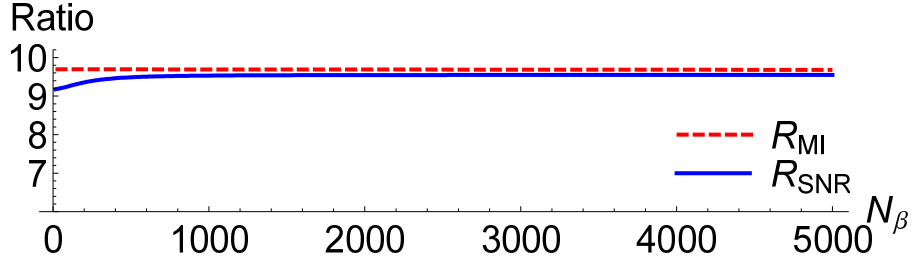
$$\hat{a}_{\text{eff}} = \frac{1}{\sqrt{M + M_\beta}} \sum_{k=1}^{M+M_\beta} \hat{a}_k, \quad (4.15)$$

where  $\hat{a}_i$  refer to the individual mode operators in our system. As before, the effective modes are therefore a linear combination of the original ones, respecting the canonical commutation relations  $[\hat{a}_{\text{eff}}, \hat{a}_{\text{eff}}^\dagger] = 1$ , and we can imagine the coarse-graining as an average of what the detector ‘sees’, since it cannot resolve individual modes.

We take the propagation of light from the source to the far-field enacted by a lens, such that we achieve a point-to-point correspondence between the modes in the source plane and the modes in the detection planes. Nevertheless, by analogy with the previous section, the results of our analysis are reproducible for near-field propagation as well [1, 110]. By merit of the far-field propagation we have  $\langle \hat{a}_k^\dagger \hat{a}_{k'} \rangle = \delta_{k,k'} \langle \hat{a}_k^\dagger \hat{a}_{k'} \rangle$  and the coarse-graining in Equation (4.15) establishes the arithmetic mean of the first order auto- and cross-correlations.

### 4.3.2 Theoretical results

Our quantities of interest will once again be the total correlations between the two effective modes, as quantified by the (Rényi) mutual information and the signal-to-noise ratio.



**Figure 4.13:** A theoretical plot of the ratios  $R_{\text{SNR}}$  (blue solid) and  $R_{\text{MI}}$  (red dashed) for quantum illumination with parameters set at realistic experimental values of  $\langle N_r \rangle = 4000$ ,  $M = 90000$ ,  $M_\beta = 50$ ,  $\eta = 0.38$  and  $\eta_\beta = 0.5$ . The asymptotic limit for  $\langle N_\beta \rangle \rightarrow \infty$  is given by Equation (4.19).

The effective two-mode covariance matrices  $\sigma_{\text{THB}, \text{TWB}}$  when the illuminated object is a balanced beam splitter and the light source S consists of THB or TWB, respectively, take the standard form

$$\sigma_S = \begin{pmatrix} a_S & 0 & c_S & 0 \\ 0 & a_S & 0 & d_S \\ c_S & 0 & b_S & 0 \\ 0 & d_S & 0 & b_S \end{pmatrix},$$

with  $S = \text{THB}, \text{TWB}$ , and:

$$a_{\text{THB}} = a_{\text{TWB}} = 1 + 2\eta\bar{n}_1, \quad (4.16a)$$

$$b_{\text{THB}} = b_{\text{TWB}} = 1 + \frac{\eta\bar{n}_1 M + 2\eta_\beta\bar{n}_\beta M_\beta}{M + M_\beta}, \quad (4.16b)$$

$$c_{\text{THB}} = d_{\text{THB}} = \eta\bar{n}_1 \sqrt{\frac{2M}{M + M_\beta}}, \quad (4.16c)$$

$$c_{\text{TWB}} = -d_{\text{TWB}} = \eta\sqrt{\bar{n}_1^2 + \bar{n}_1} \sqrt{\frac{2M}{M + M_\beta}}, \quad (4.16d)$$

where  $\bar{n}_1 = \langle N_r \rangle / (\eta M)$  is the mean photon count per mode at the source and  $\bar{n}_\beta = \langle N_\beta \rangle / (\eta_\beta M_\beta)$  is the analogous quantity with respect to values for the bath.

Keeping the SNR (4.14) as figure of merit, we can quantify the enhancement achieved by the TWB over the THB by considering the ratio of the respective SNRs [9],

$$R_{\text{SNR}} = \frac{\text{SNR}_{\text{TWB}}}{\text{SNR}_{\text{THB}}}, \quad (4.17)$$

#### 4. INTENSITY INTERFEROMETRY

---

for identical single-mode statistics. Similarly, we can analyse the ‘enhancement’ in total effective correlations by defining the corresponding ratio of the MIs,

$$R_{\text{MI}} = \frac{\mathcal{T}_{\text{TWB}}}{\mathcal{T}_{\text{THB}}} . \quad (4.18)$$

It is instructive to illustrate our findings by first plotting theoretical expectations for the comparisons of the respective ratios  $R_{\text{SNR}}$  and  $R_{\text{MI}}$ . In Figure 4.13, we keep all parameters constant apart from the bath photon-count  $\langle N_\beta \rangle$ . We notice that in the regime of highly thermal bath, which is the relevant regime for which the phenomenon of quantum illumination was defined [103, 104], the ratios of SNR and MI converge to each other and become asymptotically identical. This is observed in all useful parameter regimes. If one fixes indeed  $\langle N_\beta \rangle$  to a sufficiently large number which ensures a dominant bath (e.g.  $\langle N_\beta \rangle = 5000$ ), and lets the mean photon number  $N$  of the illuminating field vary in a broad yet realistic regime (say from  $10^2$  to  $10^4$ ) an essentially perfect identity between  $R_{\text{MI}}$  and  $R_{\text{SNR}}$  is retrieved.

As one can intuitively expect, the common value attained by both the SNR and MI ratios in practical parameter regimes ( $\langle N_\beta \rangle \gg 1$ ) for quantum illumination is determined by the cross-correlations squared, namely:

$$\lim_{\langle N_\beta \rangle \rightarrow \infty} R_{\text{SNR}} = \lim_{\langle N_\beta \rangle \rightarrow \infty} R_{\text{MI}} = \left| \frac{c_{\text{TWB}}}{c_{\text{THB}}} \right|^2 , \quad (4.19)$$

where the effective correlation elements  $c_{\text{TWB,THB}}$  are defined by Equation (4.16). The proof of this for the case of the MI ratio follows by a generalisation of the proof that the normalised intensity correlations and normalised MI correspond for the HBT effect. For what concerns the SNR ratio, when the bath is dominant, the noise terms in the denominator of the SNRs are effectively independent of the source, TWB or THB, actually considered in the protocol and cancel each other, provided they have the same single beam fluctuation (see [9] for details). Thus:

$$R_{\text{SNR}} \approx \frac{|\langle \Delta^{\text{in}} - \Delta^{\text{out}} \rangle|_{\text{TWB}}}{|\langle \Delta^{\text{in}} - \Delta^{\text{out}} \rangle|_{\text{THB}}} . \quad (4.20)$$

It can be shown through the Gaussian moment factoring that this equates the ratio of cross-correlations squared, Equation (4.19).

Interestingly, this finding provides a small extra link between the HBT effect and ghost imaging: in particular, with ghost imaging we compared the SNR to the MI

of effective operators and with the HBT effect we compared the MI to the ratio of intensity covariances [2]. Within the quantum illumination setting considered here, in which all of these quantifiers can be defined and jointly analyzed, we find that they are all quantitatively connected in the asymptotic regime of large bath.

In [9], the SNR enhancement was also linked to the ratio of the generalized Cauchy-Schwartz parameter  $\varepsilon \equiv \langle : \delta N_s \delta N_r : \rangle / (\langle : \delta^2 N_s : \rangle \langle : \delta^2 N_r : \rangle)^{1/2}$ , where  $\langle : : \rangle$  is the normally ordered quantum expectation value and  $\varepsilon \leq 1$  indicates a classical regime, i.e. corresponding to a state having positive well-defined  $P$ -representation. In particular it has been shown that the quantum enhancement with respect to the optimal classical strategy, in the limit of dominant bath, is  $R_{\text{SNR}} = \frac{\varepsilon_{\text{TWB}}}{\varepsilon_{\text{THB}}} > 1$ . Remarkably, in the presence of the bath one finds  $\varepsilon_{\text{TWB}} \leq 1$  indicating classicality although the enhancement survives. Lastly, we mention that the same transition to the classical regime without affecting the performance improvement has been observed in [134] for yet another common parameter of non-classicality, the noise reduction factor  $\langle \delta^2(N_s - N_r) \rangle / (\langle N_s + N_r \rangle)$  [132, 133, 135, 136].

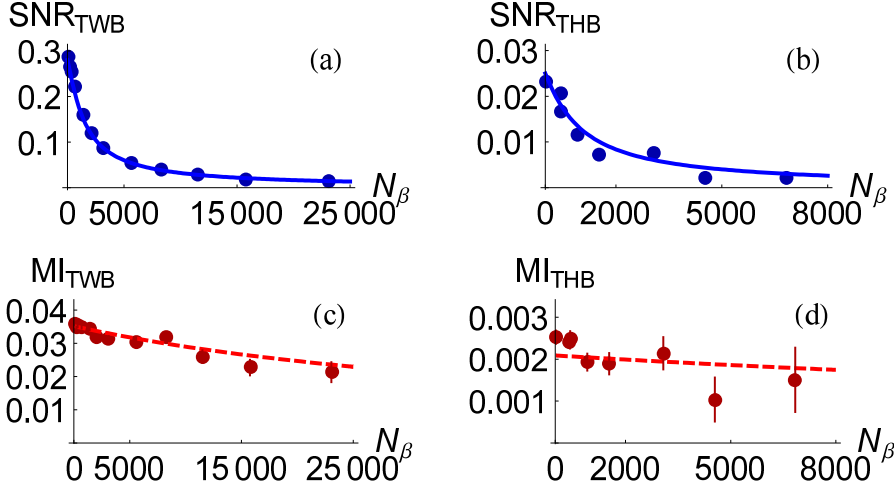
This series of observations shows how different parameters, originally introduced to assess experimental quality, are in fact all capturing the same physics in the case of Gaussian light sources for quantum illumination, and are thus all able to reveal the quantum advantage of the scheme, even in a regime in which quantumness in the form of entanglement appears not to manifestly survive.

### 4.3.3 Comparisons with experiment

The practical description introduced in the previous section allows us to directly assess the experimental results obtained in [9]. Assuming knowledge of the mode-count, we can extract our effective second-order correlations from intensity measurements and covariances. For example,  $\langle \hat{a}_{\text{eff}}^\dagger \hat{a}_{\text{eff}} \rangle = \langle N_r / M \rangle$ , from which the auto-correlations can be deduced with ease.

In Figure 4.14(a)–(b) we plot the SNRs for TWB and THB as determined experimentally in [9]. In the same figure, panels (c)–(d), we plot the corresponding MIs obtained from the experimental data by constructing the effective two-mode operators as detailed above. We observe that a very good agreement is reached with the theoretical expectations based on Eqs. (4.16), especially in the case of TWB light. The THB case is affected by considerably lower accuracy: this is consistent with the intrinsic

#### 4. INTENSITY INTERFEROMETRY



**Figure 4.14:** (a)–(d) **Experimental results and theoretical expectations for the SNRs (blue solid) and effective MIs (red dashed)** of the quantum illumination demonstration of Lopaeva *et al.* [9] using TWB and THB light, plotted versus the bath photon number  $\langle N_\beta \rangle$ . The values of the other experimental parameters are fixed at  $M = 90000$ ,  $M_\beta = 1300$ ,  $\eta = 0.38$ ,  $\eta_\beta = 0.5$ ,  $\langle N_{\text{TWB}} \rangle = 4232$ ,  $\langle N_{\text{THB}} \rangle = 3278$ , for all the plots in this figure.

lower SNR of the measurement with THB (since the enhancement ratio  $R \sim 10$ , we note that in order to achieve the same accuracy, a number of acquisitions 100 times larger than in the case of TWB illumination would be required for THB sources).

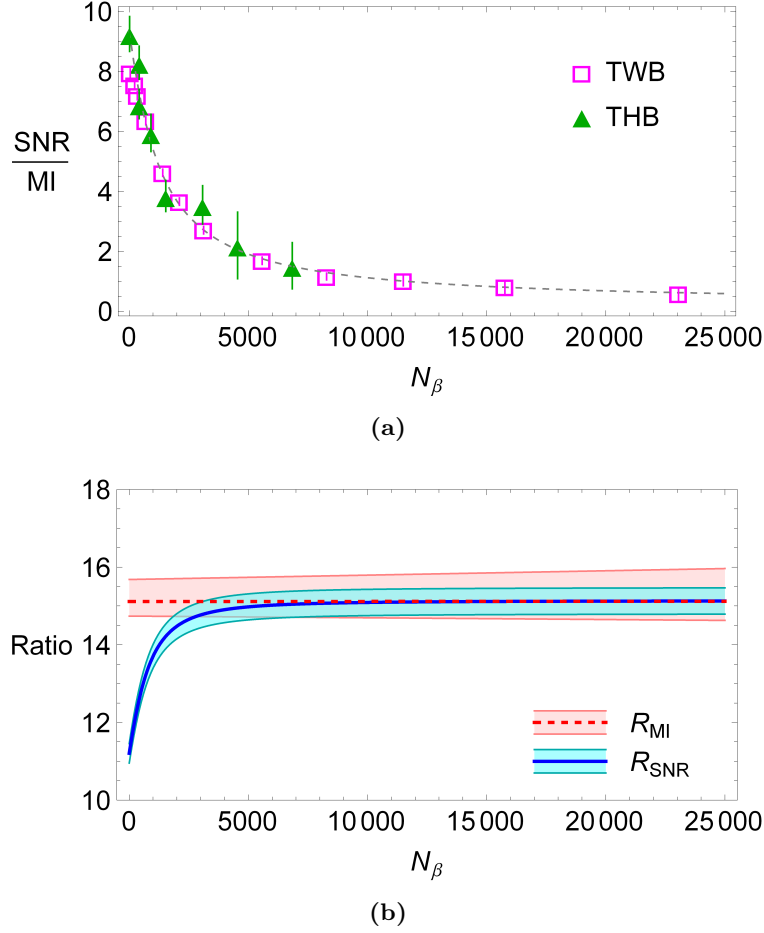
According to the theoretical expectations, we should find that

$$R_{\text{SNR}} \equiv \frac{\text{SNR}_{\text{TWB}}}{\text{SNR}_{\text{THB}}} \approx \frac{\text{MI}_{\text{TWB}}}{\text{MI}_{\text{THB}}} \equiv R_{\text{MI}},$$

or, equivalently,

$$\tilde{R}_{\text{TWB}} \equiv \frac{\text{SNR}_{\text{TWB}}}{\text{MI}_{\text{TWB}}} \approx \frac{\text{SNR}_{\text{THB}}}{\text{MI}_{\text{THB}}} \equiv \tilde{R}_{\text{THB}},$$

in the relevant regime of high  $\langle N_\beta \rangle$ . In Figure 4.15(a), we plot the ratios  $\tilde{R}$  for TWB and THB light, respectively, as calculated directly from the measured data. We see indeed that the two quantities align with good precision along the same curve, in agreement with the theory.



**Figure 4.15:** (a) Ratios  $\tilde{R}$  between SNR and MI for TWB (magenta empty squares) and THB (green filled triangles) sources, obtained from the measured data; the dashed gray curve represents the theoretical prediction, given by the mean between  $\tilde{R}_{\text{TWB}}$  and  $\tilde{R}_{\text{THB}}$ . (b) Confidence intervals (shadings) inferred from the experimental data, along with theoretical predictions (lines), for the ratios  $R_{\text{SNR}}$  (blue solid) and  $R_{\text{MI}}$  (red dashed).

Finally, to quantify the quantum enhancement in the implemented instance of continuous variable quantum illumination, we extrapolate the confidence intervals for the direct ratios  $R_{\text{SNR}}$  and  $R_{\text{MI}}$  and plot them in Figure 4.15 (b) against the theory (similarly to Figure 4.13). We conclude that the quantum enhancement allowed by the TWB over the corresponding THB with the same single-mode statistics is of a factor  $\approx 15.1$ , as determined by the asymptotic value of the ratios in the  $\langle N_\beta \rangle \gg 1$  regime, Equation (4.19). Note that we cannot directly calculate these ratios from the experimental points without resorting to model fitting, as the acquisitions in [9] correspond to

## 4. INTENSITY INTERFEROMETRY

---

different values of  $\langle N_\beta \rangle$  between the TWB and THB settings [see e.g. Figure 4.15 (e)].

### 4.4 Conclusion

We have shown that for a continuous variable realization of quantum illumination [103, 104] as demonstrated experimentally in [9], the fractional increase of mutual information (in the effective two-mode description) for entangled twin beams over correlated thermal beams provides a close approximation for the equivalent ratio of the SNR, which becomes exact in the practically relevant regime of a lossy system with a large number of thermal photons in the bath. This observation, as well as connecting to our other previous work on correlations in intensity interferometry setups, provides insight into the source of quantum improvement in continuous variable quantum illumination. We neatly observed the predicted correspondence from the experimental data of [9].

We would like to point out that these results complement and do not controvert the results of [137] or [138] for discrete variable quantum illumination. In particular, in [138] the discord consumption, i.e. the difference between the discord in the source light before and after the interaction with the target, is linked quantitatively to the quality of the protocol. Since some discord remains even when the initial entanglement is destroyed, the authors of [138] conclude that discord plays a key role in empowering quantum rather than classical illumination. In the continuous variable setting, the Gaussian discord consumption is known to relate to a quantum advantage in a simple protocol of information encoding [139], but such a scenario has not been investigated to date for the setting of quantum illumination (and it can be a good topic for further study).

We have taken a completely different approach where the correlations evaluated just for the source light are chosen as the object of study. While clearly the mutual information includes both classical and quantum portions, we find that it is the total effective correlation, rather than just the effective discord of the source, that is found to capture the quantum advantage in a quantitative fashion. The central observation is that entangled states can be overall more correlated (classically and quantumly) than separable states, for a given mean energy of the states. The resilience of these extra correlations, which we quantify via the mutual information in the effective picture, is here



shown to capture the quantum enhancement, even when external noise degrades those correlations to the point that the quantum signature of entanglement is completely suppressed.

This concludes the section on intensity interferometry. We showed for ghost imaging that although a semi-classical description is always possible, the controversy was not entirely ill-founded. It appears that the mutual information and other more general forms of correlations also feature on a deeper level in intensity interferometric schemes, as we showed experimentally with quantum illumination.

## Part III

# 5

## Multiparameter Metrology

### 5.1 Background

Quantum estimation theory was founded in the '60s and '70s, with especially notable contributions by Helstrom [28] and Holevo [29], who each produced monographs which provide excellent introductions on the subject, as well as including some novel results.

Ultimately, any estimation problem, quantum or classical, begins on the same footing. For some set of parameters described by a vector  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$  with  $\boldsymbol{\theta} \in \mathbb{R}^N$  to be estimated (*estimanda*), we define an *estimator*  $\hat{\boldsymbol{\theta}}(\mathbf{x}) = (\hat{\theta}_1(\mathbf{x}), \dots, \hat{\theta}_N(\mathbf{x}))$ , as a mapping from a  $\boldsymbol{\theta}$ -dependent M-dimensional collection of data  $\mathbf{x} = (x_1, \dots, x_M)$ , to the same space  $\mathbb{R}^N$  as the estimanda. We may choose to omit the index  $\mathbf{x}$  where it is not likely to cause confusion. We refer to the value of the estimator as the *estimate*. Obviously, this estimator shouldn't be some completely arbitrary function: we want it to provide an accurate estimate of the values of the parameter, and we thus need a means of ensuring 'closeness' of  $\hat{\boldsymbol{\theta}}$  to the true value of the parameter vector  $\boldsymbol{\theta}$ .

This is achieved by penalising deviation with the use of a *cost function*  $C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})$ . One example of a cost function is  $C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = -\delta(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$  which penalises all errors equally – or to be more precise, only rewards exact estimates. This doesn't seem like an ideal candidate when estimating continuous parameters since a tiny deviation of the estimate and a huge one would be treated on completely equal footing. A much more sensible choice, and henceforth the only cost function we shall consider, is the quadratic form

$$C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sum g_{jk}(\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k), \quad (5.1)$$

## 5. MULTIPARAMETER METROLOGY

---

where the matrix  $G = [g_{jk}]$  is positive definite. We require this of  $G$  else it would be possible to reward, rather than penalise an error. We can see that the above looks very much like a covariance matrix.

The accuracy of an estimator is then described by the expectation of the cost function, in this particular case referred to as the *total mean-square deviation*

$$\mathbb{E}[C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})] = \int C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) p(\mathbf{x}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d^n \boldsymbol{\theta} d^m \mathbf{x}, \quad (5.2)$$

where  $p(\boldsymbol{\theta})$  refers to the prior probability of finding the value  $\boldsymbol{\theta}$ , and  $p(\mathbf{x}|\boldsymbol{\theta})$  refers to the probability of finding data  $\mathbf{x}$  given the true values of the estimanda are  $\boldsymbol{\theta}$ . We have made an implicit assumption of a deterministic strategy for selecting the estimator based on a particular set of data.

When we refer to an *unbiased* estimator, we mean an estimator such that

$$\mathbb{E}_{\boldsymbol{\theta}}[\hat{\boldsymbol{\theta}}] = \int \hat{\boldsymbol{\theta}} p(\mathbf{x}|\boldsymbol{\theta}) d^m \mathbf{x} = \boldsymbol{\theta}. \quad (5.3)$$

In words, this means that at fixed  $\boldsymbol{\theta}$ , the expectation value of the estimator matches the value of  $\boldsymbol{\theta}$ . We normally relax this condition by considering *locally unbiased* estimators. An estimator is locally unbiased at  $\boldsymbol{\theta}_0$ , if the unbiasedness condition holds at that particular value but not necessarily for all values of  $\boldsymbol{\theta}$ . Additionally we require that  $\frac{d}{d\theta_j} \mathbb{E}_{\boldsymbol{\theta}}(\hat{\theta}_k) |_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \delta_{jk}$ , which can be obtained from Equation (5.3) by differentiation.

This criterion requires us to consider local estimation problems, where the prior probability distribution is very tight, meaning we already have a fairly good approximation to the parameter (formally,  $p(\boldsymbol{\theta}) = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ )<sup>1</sup>. This may seem like a stringent requirement, but an asymptotically equivalent problem can be addressed by performing some coarse estimation on a few trials of an experiment, to gain a rough estimate of  $\boldsymbol{\theta}$ , and then performing the optimal strategy for local estimation given this knowledge [140].

When considering unbiased estimates, a link to the covariance matrix of the estimator becomes clear. Defining the covariance matrix as

---

<sup>1</sup>Strictly speaking, it is often desirable to consider unbiased estimation problems with no explicit reference to a prior at all. If we denote the probability distribution of the estimator  $\hat{\boldsymbol{\theta}}$  taking a particular value  $B$  in some measurable subset of  $\mathbb{R}^N$  by  $\mu(B) = \text{Tr } \hat{\rho}_{\boldsymbol{\theta}} M(B)$  for some POVM  $\{M(B)\}$ , then the mean square deviation at a value  $\boldsymbol{\theta}$  can be written as  $\mathbb{E}_{\boldsymbol{\theta}}[C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})] = \int C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) \mu_{\boldsymbol{\theta}}(d^n \hat{\boldsymbol{\theta}})$

$$\begin{aligned}
V(\hat{\boldsymbol{\theta}}) &\equiv \mathbb{E}_{\boldsymbol{\theta}_0}[(\hat{\theta}_j - \mathbb{E}[(\hat{\theta}_j)])(\hat{\theta}_k - \mathbb{E}[(\hat{\theta}_k)])] \\
&= \int (\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k) p(\mathbf{x}|\boldsymbol{\theta}) d^m \mathbf{x},
\end{aligned} \tag{5.4}$$

it arises that the  $\mathbb{E}_{\boldsymbol{\theta}_0}[C(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})] = \text{Tr } GV(\hat{\boldsymbol{\theta}})$ , where the subscript  $\boldsymbol{\theta}_0$  implies the average is taken a particular value of  $\boldsymbol{\theta}_0$ , or considering that we have a local estimation problem, the prior is given by  $p(\boldsymbol{\theta}) = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ .

Thus far, we have not made any particular allusion to *quantum* estimation. We have merely some abstract construction where we perform the estimation on a data set  $\mathbf{x}$ , dependent on the estimanda through  $p(\mathbf{x}|\boldsymbol{\theta})$ . To begin connecting this to the quantum case, we can consider these probabilities as the probabilities of outputs of a measurement on some state  $\hat{\rho}_{\boldsymbol{\theta}}$  carrying information about the parameters to be estimated. More rigorously, we take a POVM,  $\{\Pi_{\mathbf{x}}\}$ , and can thus find a formal correspondence by setting the probabilities for outcome  $\mathbf{x}$ , through the usual quantum formalism,  $p(\mathbf{x}|\boldsymbol{\theta}) = \text{Tr } \hat{\rho}_{\boldsymbol{\theta}} \Pi_{\mathbf{x}}$ .

Still, given a particular density matrix with a particular measurement, there is nothing particularly interesting about quantum estimation. Once we have obtained a set of data through measurement, we are arguably simply performing a classical estimation problem on these probabilities.

However, quantum estimation becomes a wonderful theory in its own right, perhaps even its own little branch of statistics, when we choose to follow the popular attitude in quantum information that a density matrix  $\hat{\rho}$  can be treated as a generalisation of of a classical probability distribution  $\{p_i\}$ . Instead, then, of considering estimation on the probability output of a state with measurement, we can try to consider the optimal estimation on a state, without a specific measurement given. Essentially, this corresponds to how well we can do with the *optimal* measurement, even though this measurement may not be immediately known.

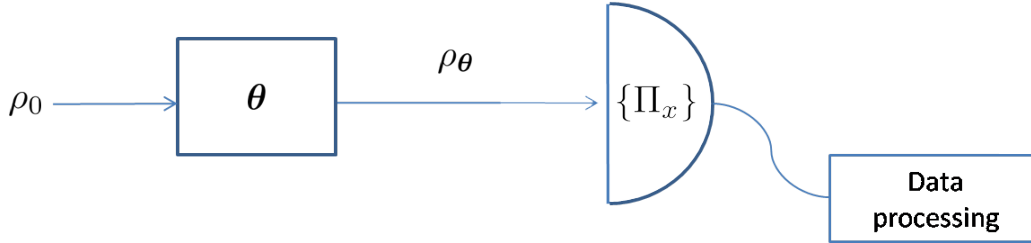
We're then presented with a wealth of complications and a correspondingly ripe field for study. This is particularly so in the case that we're estimating non-scalar  $\boldsymbol{\theta}$  (i.e. more than one parameter), where we may come to face problems of non-commutativity. An example would be if we wished to estimate the displacement parameters  $q_0$  and  $p_0$  of a vacuum state which has been acted upon by the displacement

## 5. MULTIPARAMETER METROLOGY

---

operator  $D(q_0, p_0) = e^{i(p_0\hat{q} - q_0\hat{p})}$  [141][29]. Since  $q_0$  and  $p_0$  are encoded by the non-commuting  $\hat{p}$  and  $\hat{q}$  respectively, it is not possible to directly use some of the tools of single-parameter quantum estimation.

In the quantum case, we typically consider a schematic for an estimation problem as in Figure 5.1. In addition to an optimisation over measurement, it is often desirable to consider an optimisation over the probe state  $\hat{\rho}_0$ . This is often challenging, but there exist some algorithms which simplify the task [142]. When referring to *quantum metrology*, we generally mean a system where we have control over both the quantum state and measurement and we wish to perform an estimation procedure by encoding the parameters onto the probe state.



**Figure 5.1: An abstract picture of a metrology scheme.** A probe state,  $\hat{\rho}_0$ , which may consist of many separable or entangled elements, is subjected to some process determined by the parameters  $\theta$ . The resultant state,  $\hat{\rho}_\theta$ , is measured with POVM  $\{\Pi_x\}$  and the data output is processed to produce an estimate.

### 5.1.1 Cramér-Rao bounds

One of the ultimate tools for deriving limits on estimation accuracy (assuming the above quadratic cost function), is the Fisher information. We will not present a formal derivation here, however, under the assumptions outlined in the previous section, one may define, for a conditional probability distribution of data on the estimanda, the matrix with its  $jk$  element given by

$$\text{FI}(\theta)_{jk} = \int p(\mathbf{x}|\theta) \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_j} \frac{\partial \log p(\mathbf{x}|\theta)}{\partial \theta_k} d^m \mathbf{x}. \quad (5.5)$$

This is the Fisher information matrix. Intuitively, it can be seen that it measures, in some sense, the mean sensitivity of the data to the parameter – if the derivatives

are large, the Fisher information is likewise large. The Fisher information matrix is valuable because it defines a bound on the covariance matrix of an unbiased estimator.

$$V(\hat{\boldsymbol{\theta}}) \geq \text{FI}(\boldsymbol{\theta})^{-1}. \quad (5.6)$$

This bound is always at least asymptotically achievable with the maximum likelihood estimator, which is asymptotically unbiased.

In quantum estimation there is no unique equivalent to the Fisher information. However, in the single parameter case, it is possible to define an effective analogue quantity. Intuitively, this involves replacing the probabilities in Equation 5.5 with the density matrix, and the logarithmic derivative with operator versions. For each of the estimanda  $\theta_j$ , this operator is implicitly defined by the equation

$$\frac{\partial \hat{\rho}_{\boldsymbol{\theta}}}{\partial \theta_j} = \frac{1}{2}(\hat{\rho}_{\boldsymbol{\theta}} L_{\theta_j} + L_{\theta_j} \hat{\rho}_{\boldsymbol{\theta}}), \quad (5.7)$$

$L_{\theta_j}$  is known as the symmetric logarithmic derivative (SLD). Notice how this is a symmetrised operator-version of the equation

$$\frac{\partial p(x|\boldsymbol{\theta})}{\partial \theta_j} = p(x|\boldsymbol{\theta}) \frac{\partial \log p(x|\boldsymbol{\theta})}{\partial \theta_j} = \frac{1}{2} \left( p(x|\boldsymbol{\theta}) \frac{\partial \log p(x|\boldsymbol{\theta})}{\partial \theta_j} + \frac{\partial \log p(x|\boldsymbol{\theta})}{\partial \theta_j} p(x|\boldsymbol{\theta}) \right),$$

hence the name. It is also possible to define an unsymmetrised version, usually known as the right logarithmic derivative (RLD) [29].

From this, we can define what we will call the quantum Fisher information matrix  $\text{QFI}(\hat{\rho}_{\boldsymbol{\theta}})_{jk} = \frac{1}{2} \text{Tr} \hat{\rho} \{L_j, L_k\}$ , where  $\{\cdot, \cdot\}$  is the anticommutator and the SLD Cramér-Rao bound is given by

$$V(\hat{\boldsymbol{\theta}}) \geq \text{QFI}(\boldsymbol{\theta})^{-1}. \quad (5.8)$$

exactly as in the classical case.

This Cramér-Rao bound was derived by both Holevo [29] and Helstrom [28], however, it was proven to be always achievable by Braunstein and Caves [143] when the estimation is of only a single parameter  $\theta_j$ . It is known that it is not generally achievable for multiple parameters, in which case the RLD bound may be more informative [29]. A (not necessarily unique) measurement achieving the bound for a single parameter is that where the POVM elements are the projectors forming the eigenbasis of the SLD [144].

## 5. MULTIPARAMETER METROLOGY

---

This observation immediately hints at why saturating this bound may not be possible for multiple parameters: it may be that the optimal measurements for each parameter are incompatible. This leads us to a simple condition for achievability of the SLD Cramér-Rao bound; if  $[L_j, L_k] = 0$ , then they share an eigenbasis and thus have a common optimal measurement. The question arises if, given the fact that the optimal measurements are not unique, this condition is exhaustive. It was shown by Matsumoto [145] that this is *not* the case (more precisely, there is an overlooked subtlety) and there is a more general condition for achievability of the SLD Cramér-Rao bound. This will be discussed in the next section.

It is worth considering explicit forms of the symmetric logarithmic derivatives for both pure and mixed states. For a pure state  $\hat{\rho} = |\psi\rangle\langle\psi|$ , then its SLD is given by

$$L_{\theta_j} = 2(|\psi\rangle\langle\partial_{\theta_j}\psi| + |\partial_{\theta_j}\psi\rangle\langle\psi|), \quad (5.9)$$

and for an orthogonally decomposed mixed state  $\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$

$$L_{\theta_j} = 2 \sum_{mn} \frac{\langle\psi_m|\partial_{\theta_j}\hat{\rho}|\psi_n\rangle}{p_m + p_n} |\psi_m\rangle\langle\psi_n|. \quad (5.10)$$

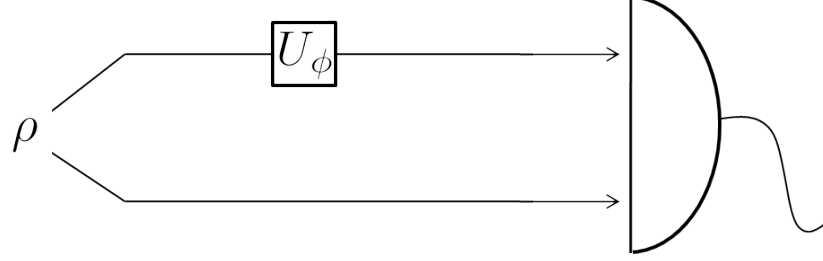
In the above equations  $|\partial_{\theta_j}\psi\rangle = \frac{\partial}{\partial\theta_j}|\psi\rangle$  and  $\partial_{\theta_j}\hat{\rho} = \sum_i (\partial_{\theta_j}p_i)(|\psi_i\rangle\langle\psi_i|) + p_i(|\psi_i\rangle\langle\partial_{\theta_j}\psi_i| + |\partial_{\theta_j}\psi_i\rangle\langle\psi_i|)$ . We have also omitted the index  $\theta$  on the quantum state. We will continue to use these conventions where it does not lead to confusion.

### 5.1.2 The Heisenberg Limit

One of the early appeals of quantum metrology was the promise of greater estimation precision using entangled probes than possible for any unentangled probes. If we consider a single-parameter estimation problem, then quantum estimation proposes that the variance of the estimator can scale as  $\text{var}(\hat{\theta}) \propto \frac{1}{N^2}$ , so-called Heisenberg scaling, whereas for separable probes the best scaling of variance is  $\text{var}(\hat{\theta}) \propto \frac{1}{N}$  which is also known as the standard quantum limit (SQL). Here  $N$  refers to the number of probes.

This is best illustrated in an interferometric scheme for phase estimation. Given  $N$  photons, we may either send them into the interferometer in a large entangled NOON state or in  $N$  separable states as in Figure (5.2). An optimal measurement strategy can be achieved by projecting the output onto the original input. For the separable states, the probability of success per trial is then given by  $\frac{1+\cos(\phi)}{2}$ , and for NOON states by





**Figure 5.2: A Mach-Zehnder interferometer for the estimation of phase.** The input,  $\hat{\rho}$ , may be chosen to be either  $N$  separable single photon states, or an entangled NOON state. The output is then either  $\frac{1}{\sqrt{2}}(|01\rangle + e^{i\phi}|10\rangle)^{\otimes N}$  or  $\frac{1}{\sqrt{2}}(|0N\rangle + e^{iN\phi}|N0\rangle)$  respectively. In the former case, we can view this as  $N$  independent trials.

$\frac{1+\cos(N\phi)}{2}$ . It is not difficult to verify that the Fisher information for these probabilities scales as  $N$  and  $N^2$  respectively. It thus appears that entanglement has great potential to reduce the covariance of the estimation for a given number of probes, via Heisenberg scaling.

Heisenberg scaling is an interesting point of appeal, but the above conception, though common, is very faulty [16]. To begin, the asymptotic attainability of the Cramér-Rao bound is based upon a large number of *independent* trials. Having a large entangled  $N$ -particle state is not equivalent to this. To more properly formulate the bound, we should consider a division of the  $N$  probe states into  $\mu \gg 1$  groups of  $n$  probes. Then as  $\mu \rightarrow \infty$ , we can apply the usual results for attainability of the Cramér-Rao bound. Thus, the inverse square enhancement is not in the total number of probes, but rather only a small fraction  $n$  corresponding to the size of the entangled states. We then get  $\text{var}(\theta) \geq \frac{1}{\mu n^2}$ .<sup>1</sup>

Moreover, it has been shown the Heisenberg limit is extremely fragile even if we do take it naively. A powerful framework for deriving ultimate bounds in the presence of noise has shown that almost universally the best improvement that can be hoped for in the scaling of the QFI with  $N$  is a constant factor improvement [147]. For the NOON

<sup>1</sup>Although Heisenberg scaling normally doesn't make sense if we use the techniques of locally unbiased estimation when  $\mu = 1$ , Hayashi [146] showed using a min-max optimisation method that for asymptotically large  $N$  it *is* possible to achieve an  $N^{-2}$  scaling in a Mach-Zehnder type setup, although with a constant factor larger than 1.

## 5. MULTIPARAMETER METROLOGY

---

state, the fragility is plain to see: if even a single photon is lost, then the superposition is completely broken as the arm the photons travelled through can be determined. We are left with a mixed state of  $N - 1$  photons in either one arm or the other and interferometry is not possible.

Further still, in practice, achieving the Heisenberg limit in an interferometric scheme may require a hard-to-achieve prior. A NOON state is invariant under phase shifts of the order of  $\frac{2\pi}{N}$ , thus we must know the phase to within an order of  $\frac{1}{N}$  before we even begin the estimation.

This does not mean that sub shot-noise scaling is not possible, but rather that one must be careful to avoid excessive optimism about the extent to which it can be beaten. For small entangled states, a constant factor improvement can be achieved over the standard quantum limit. Continuous variables have also shown significant promise, due to their easy controllability and robustness against noise.

Even so, one can legitimately ask if there is ever really a practical advantage to creating these challenging entangled states; photons, after all, are not scarce. One way to answer this question in the affirmative is to realise that it may be that the limiting resource is *not* the probe state or its components. One can conceive of a fragile test sample being a resource in of itself – the limit is then on the number of passes through the quantum channel represented by the sample. For example, in a situation involving testing of sensitive biological tissue, the scarcity may not be of photons, but of biological samples. Much like any other resource, these are expendable and their scarcity provides them with sufficient value that we must invest an extra cost elsewhere, in creating entangled states for example, to compensate.

As a final comment, Heisenberg scaling even further loses some of its novelty when we formally consider the limiting resource as the total amount of energy which can pass through a channel. Reconsidering the earlier interferometric setting, one may take a single photon state and apply the phase unitary  $N$  times. The final state will then be  $\frac{1}{2}(|01\rangle + e^{iN\phi}|10\rangle)$  and the measurement probability distribution identical to that of a NOON state, for the same total amount of energy passed. A formal mathematical equivalence between entanglement-enhanced estimation and this *sequential* estimation has been expounded upon by Maccone [148].

## 5.2 Simplifying Multiparameter Estimation

The aim of this section is to identify the circumstances under which quantum multiparameter estimation becomes almost completely analogous to quantum single parameter estimation. The first step to this is to identify the conditions under which the SLD Cramér-Rao bound is achievable (in which case we need not consider the RLD Cramér-Rao bound).

An important result on this was achieved for pure states with unitarily encoded parameters by Matsumoto [145]. Since a pure state corresponds to a rank 1 density matrix, operations acting upon it are only specified on the support of  $\hat{\rho}$  and addition of any component outside of this leaves the operator invariant. This implies that many different unitaries can produce the same final state when acting upon a pure state, for example  $U_1(\theta) = \exp(-i\frac{\theta}{2}(\sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z))$  and  $U_2(\theta) = \exp(-i\theta(\sigma_z \otimes \mathbb{1}))$  yield an identical output when acted upon the bell state  $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ , even though one operations acts on both qubits, and the other only on a single qubit.

It is thus to be expected that the SLDs too will be non-unique. Indeed, if *any* two SLDs commute, then they have a common eigenbasis and the corresponding estimanda share the same optimal measurement. Regardless of the choice of SLDs though, the QFI matrix is uniquely defined and we can thus hope that this allows for a unique way of proving achievability of the bound without checking every SLD. Fortunately, Matsumoto showed that for pure states the important quantities are vectors  $|l_{\theta_i}\rangle = L_{\theta_i}|\psi\rangle$  which *are* uniquely defined. Using any two SLDs  $L_{\theta_j}, L_{\theta_k}$ , calculated on a state  $\hat{\rho} = |\psi\rangle\langle\psi|$  the SLD Cramér-Rao bound for the two corresponding parameters can be saturated if and only if

$$\text{Im}\langle l_{\theta_j} | l_{\theta_k} \rangle = \langle \psi | (L_{\theta_j} L_{\theta_k} - L_{\theta_k} L_{\theta_j}) | \psi \rangle = 0. \quad (5.11)$$

If we take some multiparameter unitary  $U(\boldsymbol{\theta})$  acting on a probe state  $|\psi_0\rangle$  such that  $|\psi\rangle = U(\boldsymbol{\theta})|\psi_0\rangle$  with generators,  $H_{\theta_j}, H_{\theta_k}$ , then the SLD commutation condition is equivalent to the condition  $\langle \psi | [H_{\theta_j}, H_{\theta_k}] | \psi \rangle = 0$ . In other words, it is possible to choose two commuting generators  $[H'_{\theta_j}, H'_{\theta_k}] = 0$ , for which  $|\psi\rangle = e^{i\theta_j H'_{\theta_j}} e^{i\theta_k H'_{\theta_k}} |\psi_0\rangle$ . Of course, the actions described by  $U$  and  $e^{i\theta_j H'_{\theta_j}} e^{i\theta_k H'_{\theta_k}}$ , will not be equivalent on *all* states, but potentially only on the chosen probe  $|\psi_0\rangle$ .

## 5. MULTIPARAMETER METROLOGY

---

Assuming the commutation condition holds for pure states, then for estimating  $N$  parameters, there will always be a POVM with  $N + 2$  elements. However, it may be that this is an entangled measurement [149]. This differs from the case of single parameter estimation, where the optimal bound is always achievable using separable measurements [150].

It was observed by Mădălin Guță [151] that the commutation condition is also applicable to mixed states in the asymptotic limit by using techniques from local asymptotic normality (QLAN) [152]. In a rough sense, QLAN is like an enhanced version of the central limit theorem. Despite the mixed state commutation condition being informally known and used in a number of publications, to our knowledge, there is no rigorous, formal proof for it available in the literature. There are also some minor inconsistencies in its usage between publications, such as differing on whether it is merely a sufficient condition [153], or both necessary and sufficient [154].

To solve this issue, I present a formal proof which almost immediately follows from the results on QLAN, particularly those of Yamagata et al. [155]. A detailed exposition of QLAN is not required and is not included as it is beyond the mathematical scope of this thesis and its author. Simply, a brief description of their result and how it links to the commutation condition will be provided.

Ultimately, the importance of QLAN lies in its assertion that for collective measurements performed on  $n$  identically prepared systems  $\hat{\rho}^{\otimes n}$ , the optimal estimation is effectively equivalent to a that of a Gaussian shift model as  $n \rightarrow \infty$ . To link this to the commutation condition it is only necessary to show that the Holevo bound is equivalent to the SLD bound if and only if  $\text{Tr}[\hat{\rho}[L_{\theta_j}, L_{\theta_k}] = 0$ .

### 5.2.1 Holevo Cramér-Rao bound

Holevo derived a bound which is stronger than both the SLD bound and the RLD bound. Importantly, this bound is shown to be *the* optimal bound for estimation on quantum Gaussian shift models which explains its relevance when QLAN results apply. Yamagata et al. [155] provide an explicit and general construction linking asymptotic estimation to the Holevo bound via QLAN. In this section, Holevo's derivation is reproduced in order to prove that the commutation condition above is necessary and sufficient for mixed states (we assume a nonpathological problem which satisfies the mild smoothness conditions of [155], thus implying QLAN).

## 5.2 Simplifying Multiparameter Estimation

---

This section closely follows chapter VI part 7 of Holevo's book [29] and contains some formalism not present anywhere else in this thesis and is consequently quite information dense. However, this section is not required for an understanding of the remainder of this work as it serves only to provide a technical proof of the commutation criterion. It is highly recommended that the interested reader refer to Holevo's book for the full background details, as space constraints prevent a complete exposition of the topic.

To begin we introduce two pre-inner products on the space of bounded Hermitian operators  $\mathcal{B}_h(\mathcal{H})$ . For  $X, Y \in \mathcal{B}_h(\mathcal{H})$ ,

$$\begin{aligned}\langle X, Y \rangle_{\hat{\rho}} &= \text{Tr } \hat{\rho}(XY + YX) \quad \text{and} \\ [X, Y]_{\hat{\rho}} &= \text{Tr } \hat{\rho}(XY - YX).\end{aligned}\tag{5.12}$$

A pre-inner product possesses all the usual characteristics of an inner product, apart from the possibility that  $\langle X, X \rangle_s = 0$  for non-zero  $X$ . The completion of a space with respect to a pre-inner product is the completion of the quotient space over the kernel of the pre-inner product. We denote the completion of  $\mathcal{B}_h(\mathcal{H})$  with respect to  $\langle \cdot, \cdot \rangle_s$  by  $\mathcal{L}_h(\hat{\rho})$ . In the case of finite dimensions  $\mathcal{L}_h(\hat{\rho}) = \mathcal{B}_h(\mathcal{H}) / \ker \hat{\rho}$ .

We further consider the *complexification* of  $\mathcal{L}_h(\hat{\rho})$  which gives the Hilbert space,  $\mathcal{L}(\hat{\rho}) = \mathcal{L}_h(\hat{\rho}) \oplus i\mathcal{L}_h(\hat{\rho})$ . On  $\mathcal{L}(\hat{\rho})$ , we can define the commutation operator  $\mathcal{D}_{\hat{\rho}}$  by

$$\langle X, \mathcal{D}Y \rangle_{\hat{\rho}} = [X, Y]_{\hat{\rho}}.\tag{5.13}$$

For conciseness of notation, consider the convention mentioned in the footnote on p. 99 where  $\mu(B) = \text{Tr } \hat{\rho}M(B)$  denotes the probability of finding  $\hat{\boldsymbol{\theta}} \in B$  where  $B$  is a subset of some measurable space. Using these tools Holevo showed that the covariance matrix of the corresponding POVM  $\mathbf{M} = \{M(d^n\hat{\boldsymbol{\theta}})\}$  obeys the following bound:

$$V\{\mathbf{M}\} \geq [\langle X_j, X_k \rangle_{\hat{\rho}}] \pm \frac{1}{2}i[[X_j, X_k]_{\hat{\rho}}],\tag{5.14}$$

where  $X_j = \int (\hat{\theta}_j - \theta_j)M(d\hat{\boldsymbol{\theta}})$ . Rearranging this equation by setting  $\kappa_{jk} = V\{\mathbf{M}\}_{jk} - \langle X_j, X_k \rangle$ , we write the mean square error as

$$\text{Tr } GV(\mathbf{M}) = \sum_{jk} g_{jk}(\kappa_{jk} + \langle X_j, X_k \rangle).\tag{5.15}$$

## 5. MULTIPARAMETER METROLOGY

---

Since both  $\kappa_{jk}$  and  $\langle X_j, X_k \rangle$  depend on the measurement, we proceed minimising each term subject to the requirement of local unbiasedness and Equation 5.14. Only when the same measurement corresponds to both minimisations would this bound be possible to achieve.

To continue with the proof it is necessary to rewrite the above equation by using the following lemma of Holevo:

**Lemma 1.** *The elements  $X_j \in \mathcal{L}_h^2(\hat{\rho})$  and the real symmetric matrix  $\kappa_{jk}$  satisfy Equation 5.15 if and only if there exist  $Y_j \in \mathcal{L}_h^2(\hat{\rho})$  and a bounded real symmetric operator  $\mathcal{B} \in \mathcal{L}_h^2(\hat{\rho})$  such that*

1.  $Y_j = \mathcal{B}X_j$
2.  $\kappa_{jk} = \langle Y_j, \mathcal{B}(\mathbb{1} - \mathcal{B})Y_k \rangle$
3.  $\mathcal{B}(\mathbb{1} - \mathcal{B}) \geq \pm \frac{i}{2} \mathcal{B} \mathcal{D} \mathcal{B}$

This further implies that  $0 \leq \mathcal{B} \leq \mathbb{1}$  and also, we may define  $\mathcal{B}$  such that it is only non-zero on the subspace defined by  $\mathcal{L} = \text{span}\{L_i\}$ .

Skipping the details, this eventually leads to the bound

$$\text{Tr } GV(\{\mathbf{M}\}) \geq \inf \text{Tr } GB^{-1}, \quad (5.16)$$

where  $B = [\langle L_j, \mathcal{B}L_k \rangle]$  and the infimum is taken over all bounded real symmetric operators  $\mathcal{B}$ . In order to show the necessity of the commutation condition, we note that in general  $[\langle L_j, \mathcal{B}L_k \rangle] \leq [\langle L_j, L_k \rangle] = J$ . Only when this condition is achieved are the Holevo and SLD Cramér-Rao bounds the same.

For this condition to be achieved, we require that  $\mathcal{B}$  act as the identity on the space of SLDs. Given that it is only defined on this space, this implies that  $\mathcal{B}$  is the projector onto this space. Invoking criterion 3 of Lemma 1, we can write

$$0 = \langle L_j, \mathcal{B}(\mathbb{1} - \mathcal{B})L_k \rangle \geq \pm \frac{i}{2} \langle L_j, \mathcal{B} \mathcal{D} \mathcal{B} L_k \rangle. \quad (5.17)$$

However, since  $\mathcal{B}$  is the projector onto the space of SLDs  $\langle L_j, \mathcal{B} \mathcal{D} \mathcal{B} L_k \rangle = \langle L_j, \mathcal{D} L_k \rangle = \text{Tr } \hat{\rho}(L_j L_k - L_k L_j)$  and thus the commutation criterion is a necessary and sufficient criterion for achieving the SLD Cramér-Rao bound.

## 5.2 Simplifying Multiparameter Estimation

The Holevo bound can be written in the somewhat more explicit form of [156]

$$\inf \operatorname{Tr} GB^{-1} = \min_{B,H} \{ \operatorname{Tr} GB^{-1}; B^{-1} \text{ is a real matrix such that } B^{-1} \geq H, \\ H_{jk} = \operatorname{Tr} \hat{\rho} H_k H_j, \text{ and } H_i \text{ satisfy the unbiasedness condition} \}$$

This allows us to explicitly construct the Hermitian operators  $H_1, H_2$  which achieve the SLD bound assuming the commutation condition holds.

Consider the QFI matrix  $J = \langle L_j, L_k \rangle$  and for simplicity restrict consideration to two variables, and let  $\begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = J^{-1} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$ . Further for some matrix  $\mathbf{A} = [A_{jk}]$ , let  $\operatorname{Tr}' \hat{\rho} \mathbf{A} = [\operatorname{Tr} \hat{\rho} A_{jk}]$ , then,

$$\begin{aligned} H &= \operatorname{Tr}' \hat{\rho} \left( \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} \cdot \begin{pmatrix} H_1 & H_2 \end{pmatrix} \right) \\ &= \operatorname{Tr}' \hat{\rho} \left( J^{-1} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} \cdot \begin{pmatrix} L_1 & L_2 \end{pmatrix} (J^{-1})^T \right) \\ &= J^{-1} \begin{pmatrix} \operatorname{Tr} \hat{\rho} L_1^2 & \operatorname{Tr} \hat{\rho} L_1 L_2 \\ \operatorname{Tr} \hat{\rho} L_2 L_1 & \operatorname{Tr} \hat{\rho} L_2^2 \end{pmatrix} (J^{-1})^T \\ &= J^{-1}. \end{aligned}$$

The final step was obtained by noting that if  $\operatorname{Tr} \hat{\rho} [L_1, L_2] = 0$ , then the middle matrix is simply QFI matrix itself. This thus indicates that the Holevo Cramér-Rao bound is equivalent to the inverse of the QFI matrix i.e. the SLD Cramér-Rao bound.

We can now relieve ourselves from the preceding abstraction. In the remainder of the thesis we shall only consider cases in which the Holevo and SLD Cramér-Rao bounds coincide. Moreover, we will only look at estimation schemes with finite dimensional inputs.

### 5.2.2 Diagonal quantum Fisher information

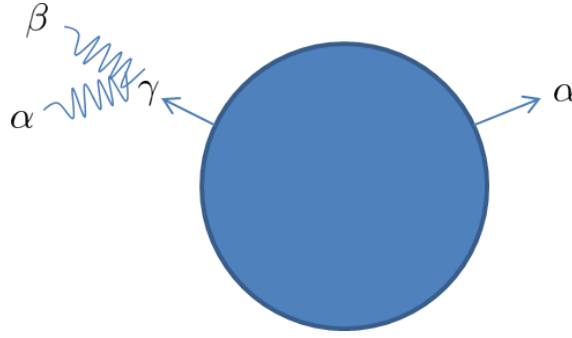
We have now established a condition on the achievability of the quantum multiparameter Cramér-Rao bound. If this condition is met, then – in a sense – we can treat our problem as a classical estimation problem with a set probability distribution. We

## 5. MULTIPARAMETER METROLOGY

---

need not concern ourselves excessively with problems of incompatible measurements and non-commutativity.

We next turn our attention to an issue which also arises in classical estimation. If the Fisher information matrix is not diagonal, then it is not possible to *independently* estimate the parameters, that is the estimators cannot be calculated separately. For example, we may be able to estimate the sum of parameters, but not their difference which would potentially render a metrological scenario useless. Intuitively, there is the simple concern of propagation of errors; if the Fisher information matrix is not diagonal the covariance matrix will also not be diagonal and this correlation implies correlated errors (Figure 5.3). We will present a more striking example of how this enters into quantum estimation later.



**Figure 5.3: A simple example of an estimation process in which we expect to have non-diagonal Fisher information terms present.** Take a source which emits two different particles, one long-lived  $\alpha$  and the other a short-lived  $\gamma$  which quickly decays into long-lived  $\alpha$  and  $\beta$ . Take the emissions as normally-distributed random processes dependent on some rates. We cannot directly observe these rates, but must deduce them from some counting statistics. Let's assume that  $\gamma$  is too short lived to directly observe, but  $\alpha$  and  $\beta$  are countable, thus the rate of the  $\alpha$ -only production process is only deducible by subtracting the  $\beta$  count from the  $\alpha$  count. This means the uncertainty in the  $\beta$  count will propagate into the estimate of the  $\alpha$ -only process, and we can expect a covariance matrix of  $\begin{pmatrix} \sigma_\beta^2 & -\sigma_\beta^2 \\ -\sigma_\beta^2 & \sigma_\beta^2 + \sigma_\alpha^2 \end{pmatrix}$ , where  $\sigma_{\alpha,\beta}^2$  are the respective variances of the  $\gamma$  and  $\alpha$ -only processes. Working backwards, this of course also entails non-diagonal Fisher information.

The condition for having a 0 diagonal term is  $\text{Tr } \hat{\rho}(L_j L_k + L_k L_j) = 0$ . If we combine this with the commutation condition, then we find that *both* conditions are satisfied when  $\text{Tr } \hat{\rho}(L_j L_k) = 0$ . It is interesting to consider the necessary conditions for satisfying this criterion and this shall be done in the following section.



In practical terms, if we have a diagonal Fisher information matrix, then this also simplifies the choice of weight matrix  $G$  significantly as we need only focus on diagonal matrices. This amounts to a simple weighting of the variances. In the quantum case, this will affect the choice of probe state, so for two parameters we would minimise  $a \text{var}_{\hat{\rho}}(\hat{\theta}_1) + (1-a) \text{var}_{\hat{\rho}}(\hat{\theta}_2)$  over a choice of input state  $\hat{\rho}_0$  which is evolved into  $\hat{\rho}$  by the dynamics. Here  $0 \leq a \leq 1$  is free to be chosen according to the importance we assign each parameter.

We thus can say that two parameters are simultaneously, independently and *optimally* estimable if it is the case that the choice of  $\hat{\rho}_0$  is independent of  $a$ . In other words, the state which minimises the individual variances is the same for both parameters.

To summarise the language we use, simultaneous estimability shall refer to satisfaction of the commutation condition and thus the ability to saturate the Cramér-Rao bound. Simultaneous and *independent* estimability refers to the additional satisfaction of the anticommutation condition, or in the language of [145], ‘informational independence’ of the parameters. Lastly, recalling that we’re considering a metrological scenario, if the optimal probe state is also the same for both parameters, then we say the parameters are *optimally* simultaneously and independently estimable.

## 5.3 Application

### 5.3.1 Preliminary Observations

Explicit formulae for the SLDs were given in Section 5.1.1. For convenience, we repeat them here and provide some additional insights. To set a standard notation, we henceforth denote  $\frac{\partial}{\partial x}$  by  $\partial_x$ .

For a mixed state, and some parameter  $\theta_i$ , the SLD is given by

$$L_{\theta_i} = \sum_{m,n} \frac{\langle \psi_m | (\partial_{\theta_i} \hat{\rho}) | \psi_n \rangle}{p_m + p_n} |\psi_m\rangle \langle \psi_n|. \quad (5.18)$$

Accordingly, we can calculate the term,

$$\text{tr} \hat{\rho} L_{\theta_i} L_{\theta_j} = \sum_{m,n} \frac{p_m}{(p_m + p_n)^2} (\langle \psi_m | \partial_{\theta_i} \hat{\rho} | \psi_n \rangle \langle \psi_n | \partial_{\theta_j} \hat{\rho} | \psi_m \rangle), \quad (5.19)$$

## 5. MULTIPARAMETER METROLOGY

---

which allows us to deduce the quantum Fisher information matrix  $\text{QFI}_{ij} = [\text{Re}(\text{Tr } \hat{\rho} L_{\theta_i} L_{\theta_j})]$ , as well as the commutation condition which may be written as  $\text{Im}(\text{Tr } \hat{\rho} L_{\theta_i} L_{\theta_j}) = 0$ .

It is particularly interesting to consider when the parameters are simultaneously and independently estimable i.e. when  $\text{Tr } \hat{\rho} L_{\theta_i} L_{\theta_j} = 0$ . One simple case where this occurs is when one parameter is encoded on the density matrix by a unitary, and the other is quasi-classical. By a quasi-classical parameter  $\theta_i$ , we mean that for  $\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ , it turns out that  $\partial_{\theta_i} \hat{\rho} = \sum_k \frac{\partial p_k}{\partial \theta_i} |\psi_k\rangle\langle\psi_k|$ . This resembles a classical probability distribution when represented by a diagonal matrix.

The unitary encoding of the parameter  $\theta_j$  will render all terms of the form  $\langle\psi_n|\partial_{\theta_j}\hat{\rho}|\psi_n\rangle$  equal to 0. On the other hand, for  $\langle\psi_n|\partial_{\theta_i}\hat{\rho}|\psi_m\rangle$ , a quasi-classical  $\theta_i$ , ensures these terms are 0 for  $m \neq n$ . Combined, these two conditions imply that the numerator of every term of the sum in Equation (5.19) is 0. It is also possible for each term of the numerator to be 0, but with a more complex interplay between the parameters, as will be shown in the section on simultaneous estimation of phase and dephasing.

One may imagine that there are yet more complicated ways of the sum equalling 0, involving cancellations between terms, rather than each term on its own equalling 0. It is worth asking if some general, easily verifiable conditions for quantum channels being simultaneously and independently estimable can be found but the problem is not explored here.

For pure states, with unitarily encoded parameters, the formulae are much simpler. The SLD is provided by  $L_{\theta_i} = 2(|\partial_{\theta_i}\psi\rangle\langle\psi| + |\psi\rangle\langle\partial_{\theta_i}\psi|)$ . One then finds,

$$\text{Tr } \hat{\rho} L_{\theta_i} L_{\theta_j} = \langle\psi|H_{\theta_i}H_{\theta_j}|\psi\rangle - \langle\psi|H_{\theta_i}|\psi\rangle\langle\psi|H_{\theta_j}|\psi\rangle,$$

where  $H_{\theta_i}$ ,  $H_{\theta_j}$  are the generators of the unitaries which encode the parameters  $\theta_i$  and  $\theta_j$  respectively.

### 5.3.2 Estimation of two unitaries on qubits

The estimation of multiple unitary parameters has been considered before in [149, 153, 157]. In [149, 157] qubit unitaries of the form  $U(\alpha, \theta, \phi) = \cos \alpha \mathbb{1} + \sin \alpha \vec{n} \cdot \vec{\sigma}$  where  $\vec{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^T$  and  $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$  were studied. It was found that

by entangling the qubit with an ancillary system such that the estimation takes place on  $U(\alpha, \theta, \phi) \otimes \mathbb{1} |\psi_0\rangle$  for a maximally entangled 2-qubit input state  $|\psi_0\rangle$ , it was possible to achieve the Cramér-Rao bound simultaneously for all 3 parameters. In [153], they considered two-phase spin rotations on systems of arbitrary spin degree and found that simultaneous estimability was possible and approximate scaling of variances of roughly  $\frac{1}{j^2}$  could be achieved in the estimators for each of the parameters, where  $j$  was the spin of the input.

We will consider a different two-parameter class of unitary operations in order to produce a greater enhancement in quantum multiparameter metrology than has been observed in any previous work. Below we elaborate a model of a unitary operation composed of two sequential unitaries acting on qubits. We find that for single qubit inputs it is not generally possible to simultaneously estimate both unitary phase rotations: the best strategy available amounts to only getting information about one phase at a time. However, with entanglement, and in the case that the two unitaries enact rotations around orthogonal axes of the Bloch sphere, we can simultaneously and independently estimate their phases, and particularly with two-qubit entangled states, we can obtain optimal Heisenberg-limited scaling in both parameters simultaneously.

To motivate this, we choose a physical setup consisting of two polarisation-rotating waveplates (Figure (5.4)) of known orientations but unknown thicknesses and ask the question: How best can we estimate the thicknesses of the waveplates?

It has been known since the 19th century that a birefringent crystal acts as a polarisation rotator for light. Due to the different indices of refraction along orthogonal axes, the corresponding polarisation components will propagate at different speeds, which induces a relative phase described by the equation,

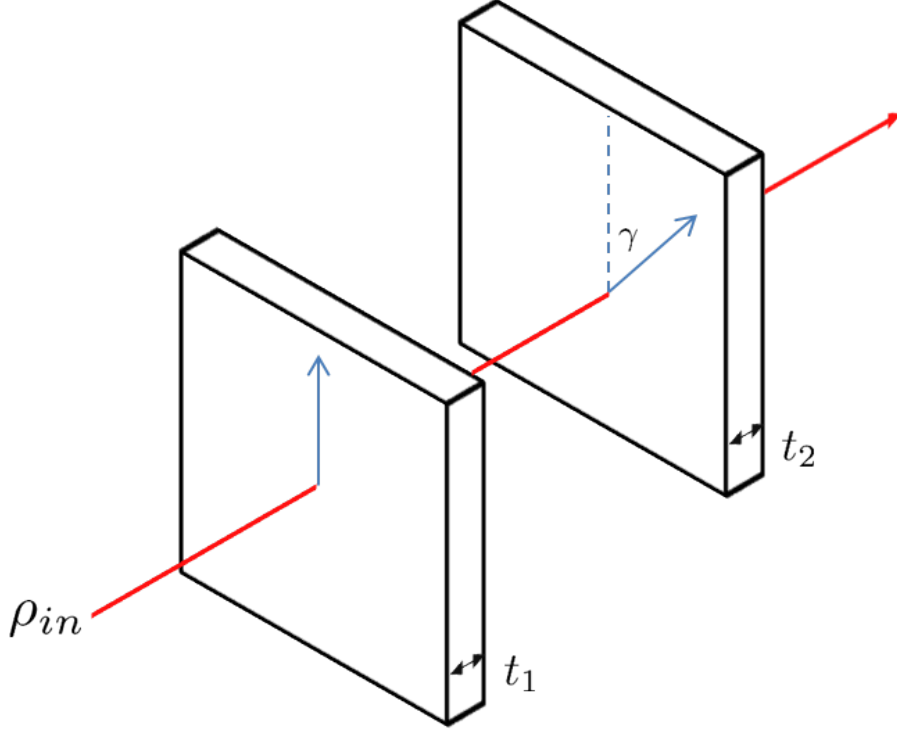
$$\Delta\phi = \frac{2\pi x \Delta n}{\lambda_0}, \quad (5.20)$$

where  $x$  is the propagation distance through the birefringent crystal,  $\Delta n$  is the difference between refractive indices for orthogonal polarisation and  $\lambda_0$  is the vacuum wavelength of the light.

Using the Jones parametrisation of polarised light, we find an effectively identical formalism to that of the qubit picture. We can set vertical polarisation  $|V\rangle \equiv |1\rangle$  and horizontal polarisation  $|H\rangle \equiv |0\rangle$ . A waveplate with the optical axis at  $\frac{\gamma}{2}$  degrees to the vertical is then described by the unitary operator  $e^{-i\frac{\gamma}{2}\sigma_x} e^{-i\frac{\Delta\phi}{2}\sigma_z} e^{i\frac{\gamma}{2}\sigma_x}$ .

## 5. MULTIPARAMETER METROLOGY

---



**Figure 5.4: Sequential waveplates.** The first plate has its optical axis aligned vertically, and the second at  $\gamma$  degrees to the vertical. Their thicknesses are  $t_1$  and  $t_2$  respectively.

In estimating the thicknesses of the waveplates, assuming we avoid multi-pass experiments, there are two ways we can conceive of performing the experiment with finite resources ( $N$  photons). Assuming we are equally interested in both thicknesses, we can divide the photons into two groups of  $\frac{N}{2}$  and perform an optimal estimation on each. This will lead to scalings of  $\text{var}(\hat{\theta}) \propto \frac{4}{N^2}$  each (keeping in mind the earlier caveats about naive treatment of Heisenberg-limit scaling). More interesting would be to attempt to use all  $N$  photons to probe both parameters. This could be done by putting the two waveplates in sequence and passing some  $N$  photon state through both of them.

We can actually use the scheme of two sequential waveplates to model *any* unitary consisting of two independent, sequential unitary operations on a qubit. That is, if we have some unitary  $U \equiv U_\theta U_\phi = e^{-i\frac{\theta}{2}\mathbf{u}_2 \cdot \boldsymbol{\sigma}} e^{-i\frac{\phi}{2}\mathbf{u}_1 \cdot \boldsymbol{\sigma}}$ , then we can without loss of generality set  $\mathbf{u}_1$  and  $\mathbf{u}_2$  such that  $U = e^{-i\frac{\theta}{2}(\cos(\gamma)\sigma_z + \sin(\gamma)\sigma_x)} e^{-i\frac{\phi}{2}\sigma_z}$ . Due to the symmetries of the Bloch sphere, this construction characterises the problem fully.

Taking a pure input state  $|\psi_0\rangle$ , and an appropriate form of the pure state form of the commutation condition, we find,

$$\begin{aligned}\text{Tr}(\hat{\rho}[L_\phi, L_\theta]) &= 4(\langle\partial_\phi\psi|\partial_\theta\psi\rangle - \langle\partial_\theta\psi|\partial_\phi\psi\rangle) \\ &= \langle\psi_0|[\sigma_z, U_\phi^\dagger(\cos(\gamma)\sigma_z + \sin(\gamma)\sigma_x)U_\phi]|\psi_0\rangle \\ &= \langle\psi_0|[\sigma_z, \sin(\gamma)(\cos(\phi)\sigma_x + \sin(\phi)\sigma_y)]|\psi_0\rangle.\end{aligned}\quad (5.21)$$

Thus the dependence of the estimation on the probe state comes from terms of the form  $\langle\psi_0|\sigma_x|\psi_0\rangle$  and  $\langle\psi_0|\sigma_y|\psi_0\rangle$ . If we choose a generic single qubit pure state  $|\psi_0\rangle = \begin{pmatrix} e^{-i\frac{\alpha}{2}} \cos(\beta/2) \\ e^{i\frac{\alpha}{2}} \sin(\beta/2) \end{pmatrix}$ , then we ultimately come to the explicit form of the commutation criterion.

$$2i \sin(\beta) \sin(\gamma) \sin(\alpha + \phi) = 0. \quad (5.22)$$

For now set  $\gamma = \frac{\pi}{2}$  such that we are dealing with waveplates which impart rotations around orthogonal axes (i.e.  $U = e^{-i\frac{\theta}{2}\sigma_x}e^{-i\frac{\phi}{2}\sigma_z}$ ). From this, we can see that no pure state input can hope to attain the SLD Cramér-Rao bound *unless* it is in the plane containing the eigenstates of  $\sigma_z$  and  $U_\phi^\dagger\sigma_xU_\phi$ .

Intuitively, this exception is easy to understand: if the probe is an eigenstate of  $\sigma_z$ , then it is in fact pointing along the axis of rotation and thus is unaffected by the initial  $U_\phi$  component of the unitary. If the initial state is an eigenstate of  $U_\phi^\dagger\sigma_xU_\phi$  then it is rotated by  $U_\phi$  into an eigenstate of  $\sigma_x$ , and the second component of the unitary has no effect. In all other cases, the final state has been affected by both rotations. However if in the plane which satisfies the commutation condition, we now have a compromise of partial ignorance of each parameter from the preceding argument.

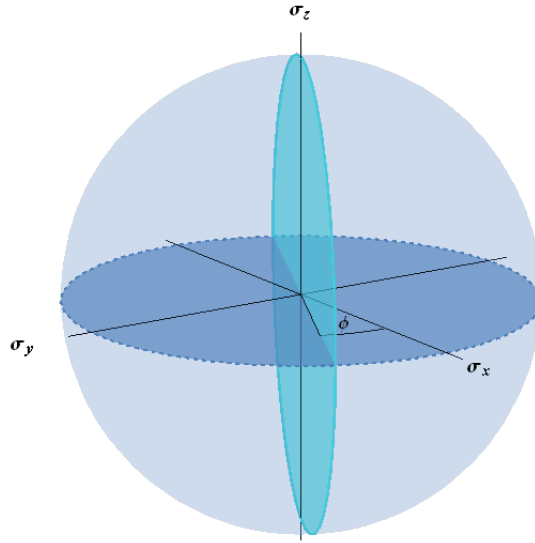
The nail in the coffin for attempting to estimate the parameters corresponding to orthogonal rotations is hammered when looking at the off-diagonals of the quantum Fisher information matrix, which we ideally wish to be 0. We find this amounts to requiring

$$\frac{1}{2} \sin(\beta)[\cos(\gamma) \sin(\beta) - \cos(\beta) \cos(\alpha + \phi) \sin(\gamma)] = 0.$$

This is only satisfied when the initial state lies in one of the two planes perpendicular to that satisfying the commutation condition. This seems to imply that if we wish

## 5. MULTIPARAMETER METROLOGY

---

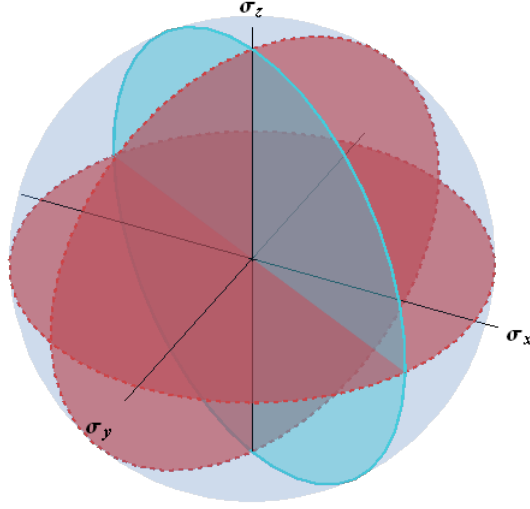


**Figure 5.5: A plot of the states which allow saturation of the Cramér-Rao bound.** The turquoise (vertical) plane contains the  $\sigma_z$  axis, and lies at angle of  $-\phi$  from the  $\sigma_x$  axis. This plane contains all of the states with satisfy the commutation condition.

to satisfy the commutation condition we will always have off-diagonal terms. Only on the intersection of the planes are both conditions satisfied, however, here we are handicapped by the incompatible input states and are doomed to gain 0 knowledge about one of the parameters. Thus, any way we attempt to estimate the parameters we *must* sacrifice knowledge of one for knowledge of another. We gain no advantage over using entirely separate probes for the estimation of each parameter.

Of course, instead of orthogonal rotation axes, we may choose the opposite extreme. By setting  $\gamma = 0$ , we have two rotations with the same generator,  $\sigma_z$ . From a glance at the commutation condition, we see that it is automatically satisfied in this situation. This should come as no surprise at all, since the optimal measurement for estimating any phase rotation lies in the plane orthogonal to the axis of rotation (the equator of the Bloch sphere, in this case), and does not depend on the value of the phase itself.

It should also come as no surprise that the diagonality condition is maximally violated. We cannot hope to distinguish the rotations by any measurement at all. While we gain full information about the sum of the two phases, we know nothing of their difference. States with  $0 < \gamma < \frac{\pi}{2}$  will always sit on some intermediate level of compromise between the two extremes.



**Figure 5.6: An expanded version of Figure 5.5.** The turquoise plane satisfies the commutation condition as before. We also plot two red planes (with dashed borders). Any states on these two planes will have a diagonal QFI matrix. Only the states on the intersections of a red plane with the turquoise plane satisfy both conditions, however, these states yield 0 information about one parameter.

When using separable photons, the situation is thus dire indeed. Can the magic of entanglement somehow overcome the numerous obstacles encountered by separable qubits?

To begin answering this question, we need only some simple identities. Assume all  $N$  probe photons are identically and independently affected by the unitary, then the effective  $N$ -dimensional unitary becomes  $U^{\otimes N}$  and by using the identities  $V_1 W_1 \otimes V_2 W_2 = (V_1 \otimes V_2)(W_1 \otimes W_2)$  and  $e^{-iX} \otimes e^{-iY} = e^{-i(X \otimes \mathbb{1} + \mathbb{1} \otimes Y)}$  we can write.

$$U^{\otimes N} = e^{-\frac{i}{2}\theta(\sum_{\pi} \pi[\sigma_d \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}])} e^{-\frac{i}{2}\phi(\sum_{\pi} \pi[\sigma_z \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}])}.$$

Where  $\sigma_d = \cos(\gamma)\sigma_z + \sin(\gamma)\sigma_x$  and  $\pi[\cdot]$  denotes the unique permutations of  $\sigma_{d,z}$  in the tensor product with  $N - 1$  identity matrices.

It turns out that we find an exactly analogous expression of Equation (5.21) but with the substitution of  $\sigma_z$  with  $\sigma_z^{(N)} \equiv \sum_{\pi} \pi[\sigma_z \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}]$  and similarly  $\sigma_e = U_{\phi}^{\dagger} \sigma_d U_{\phi}$  with  $\sigma_e^{(N)} \equiv \sum_{\pi} \pi[\sigma_e \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}]$ . However, there is nevertheless a crucial difference: we are allowed to have an entangled input state.

We consider a GHZ type state written as  $|\psi_0^{(N)}\rangle = \frac{1}{\sqrt{2}}(|\psi_0\rangle^{\otimes N} + |\psi_0^{\perp}\rangle^{\otimes N})$ , where  $|\psi_0\rangle$  is an arbitrary qubit. In the appropriate modification of Equation (5.21), we note

## 5. MULTIPARAMETER METROLOGY

---

that any terms of the form  $\langle \psi_0 |^{\otimes N} (\sigma_i \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}) | \psi_0^\perp \rangle^{\otimes N}$  (where  $\sigma_i$  is any linear combination of Pauli matrices) and its permutations, will cancel out, thereby leaving no cross-terms. Moreover, owing to the fact that  $\langle \psi_0 | \sigma_i | \psi_0 \rangle = -\langle \psi_0^\perp | \sigma_i | \psi_0^\perp \rangle$ , the other terms also cancel. This implies that the entanglement opens up simultaneous estimation to a much larger class of states than for a single qubit!

When calculating the off-diagonals, we find that

$$\text{Tr } \hat{\rho} \frac{\{L_\phi, L_\theta\}}{2} = \langle \psi_0^{(N)} | \frac{\{\sigma_z^{(N)}, \sigma_e^{(N)}\}}{2} | \psi_0^{(N)} \rangle - \langle \psi_0^{(N)} | \sigma_z^{(N)} | \psi_0^{(N)} \rangle \langle \psi_0^{(N)} | \sigma_e^{(N)} | \psi_0^{(N)} \rangle.$$

This is not state independent and its general solution is not particularly informative. Choosing  $|\psi_0^{(N)}\rangle = \frac{1}{\sqrt{2}}(|0^{\otimes N}\rangle + |1^{\otimes N}\rangle)$  we find that the second term is necessarily 0 since  $\langle \psi_0^{(N)} | \sigma_z^{(N)} | \psi_0^{(N)} \rangle = N \langle \psi_0^{(1)} | \sigma_z | \psi_0^{(1)} \rangle = 0$ . We find a similar result when  $|\psi_0^{(N)}\rangle$  is an analogous superposition of states in the eigenbasis of  $\sigma_e$ , but we shall without loss of generality restrict ourselves to the former case. In the case of  $N \geq 3$ , the remaining term then amounts to,

$$\begin{aligned} \langle \psi_0^{(N)} | \frac{\{\sigma_z^{(N)}, \sigma_e^{(N)}\}}{2} | \psi_0^{(N)} \rangle &= N \langle \psi_0^{(1)} | \frac{\{\sigma_z, \sigma_e\}}{2} | \psi_0^{(1)} \rangle \\ &\quad + N(N-1) \langle \psi_0^{(1)} | \sigma_z | \psi_0^{(1)} \rangle \langle \psi_0^{(1)} | \sigma_e | \psi_0^{(1)} \rangle. \end{aligned}$$

Following the calculation through, we conveniently find that we get a contribution of  $\langle 0 | \sigma_e | 0 \rangle = \cos(\gamma)$  in both terms for a total off-diagonal term corresponding to  $N^2 \cos(\gamma)$ . This indicates that only when the rotations are around orthogonal axes of the Bloch sphere can the parameters be estimated independently. As one would expect, two unitary rotations along the same axis remain indistinguishable.

Choosing the more interesting case of  $\gamma = 0$ , we can finally calculate the diagonal elements of the quantum Fisher information matrix.

$$\begin{pmatrix} N^2 & 0 \\ 0 & N \end{pmatrix}.$$

where the upper left corner corresponds to  $QFI_{\phi\phi}$ . It is possible to choose states such that we swap the scalings, or have some compromise. This shows that given an ideal noiseless setting we can indeed find an improvement on the best scheme involving two separate estimations. Specifically, we average a Fisher information of  $\frac{N^2+N}{2}$  rather than  $\frac{N^2}{4}$  per parameter, leading to a factor of 2 improvement for large  $N$ .



It is notable that the case of  $N = 2$  has not been included in the above calculation. The difference lies that when there are only 2 qubits, instead of finding terms of the form  $\langle \psi_0 |^{\otimes N} \sigma_i \otimes \sigma_i \otimes \mathbb{1} \otimes \dots | \psi_0^\perp \rangle^{\otimes N}$ , which necessarily cancel because of the identity matrices yielding inner products of orthogonal states, we have a non-zero contribution from, e.g.  $\langle 00 | \sigma_e \otimes \sigma_e | 11 \rangle$ .

The QFI matrix obtained from assuming orthogonal rotations, i.e.  $\sigma_d = \sigma_x$  is,

$$\begin{pmatrix} 4 & 0 \\ 0 & 4 \cos^2(\alpha + \phi) \end{pmatrix}.$$

It might seem that this once again reduces us to an average of  $N$  scaling in one of the parameters; however, we can take advantage of a trick owing to our knowledge of a tight prior to tailor the input state. By setting  $\alpha \approx -\phi$ , we achieve a value of 4 for the Fisher information of both  $\theta$  and  $\phi$ .

In the introductory chapters, it was mentioned that it is possible to find a (possibly non-separable) measurement which achieves this bound. Making the obvious guess of a measurement in the Bell basis, we find that the POVM given by

$$\begin{aligned} \Pi_1 &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)(\langle 11| + \langle 00|), \\ \Pi_2 &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)(\langle 11| - \langle 00|), \\ \Pi_3 &= \mathbb{1} - \Pi_1 - \Pi_2. \end{aligned}$$

achieves this bound.

This three element Bell-type measurement can be achieved deterministically by linear optics [158, 159]. This makes the 2 qubit case particularly interesting and worth further investigation as to experimental suitability for providing a proof-of-principle experiment for sub-shot-noise scaling in two parameters at a time.

To summarise the most interesting results of this section, we have found that with single-qubit inputs, and a pair of sequential unitaries around mutually orthogonal axes of the Bloch sphere, we can only hope to gain information about one of the parameters at a time. However, by allowing multi-qubit entangled inputs we circumvent this issue. In particular for two-qubit Bell-type inputs, we obtain Heisenberg scaling in both parameters.

## 5. MULTIPARAMETER METROLOGY

---

### 5.3.3 Estimation of phase and dephasing

Now let's turn our attention away from pure state models. Consider not just a unitary phase evolution, but also dephasing. Dephasing is a form of decoherence which can arise as a consequence of an imperfect phase-rotation (for example, if thermal fluctuations were to alter one of the waveplate thicknesses in the previous section). It can be viewed as the introduction of random phases to the system. A general dephasing process on a state  $\hat{\rho}$  of  $N$  qubits can be written as

$$\hat{\rho}_{\text{deph}} = \int d\phi_1 \dots d\phi_N p(\phi_1, \dots, \phi_N) U_1(\theta_1) \dots U_N(\theta_N) \hat{\rho} U_N^\dagger(\theta_N) \dots U_1^\dagger(\theta_1), \quad (5.23)$$

where  $U_j(\phi_j) = e^{-i\frac{\phi_j}{2}\sigma_z}$  are phase rotating unitaries acting on the  $j^{\text{th}}$  qubit. We have assumed without loss of generality that the phase rotation is around the  $\sigma_z$  axis.

There have been numerous studies on phase estimation in the presence of dephasing. These include phase estimation in the presence of collective dephasing [160], where the random phases on each qubit are fully correlated; in the presence of independent dephasing [161], where there is no correlation at all in the random phases; and generalised dephasing where partial correlations can exist [162].

There have also been studies on simultaneous estimation of both phase and collective dephasing [160], as well as phase and independent dephasing, including experiments in the latter case [154]. However, for phase and independent dephasing, the ultimate bounds have yet to be found, as [154] only studied the case of two dimensional subspaces. In the following section, the general bound is derived.

We will identify a class of multi-qubit entangled states which are known to contain the optimal probes for estimating phase in the presence of independent dephasing. We will show that in this class of states, it is always possible to simultaneously and independently estimate phase and dephasing although the optimal probe states differ. Our derivation will involve performing significant simplifications on the structure of the states by taking advantage of their large degree of symmetry, and then showing that the numerator of Equation 5.19 always comes to 0.

Before commencing with the derivation of the bound it is instructive to consider the single qubit case. The Kraus operators for dephasing are  $K_1 = \sqrt{\frac{1-\eta}{2}}\sigma_z$  and  $K_2 = \sqrt{\frac{1+\eta}{2}}\mathbb{1}$ . Here  $\eta$  is the equatorial radius after dephasing (as dephasing can be

seen as a constriction of the equator of the Bloch sphere). When  $\eta = 1$ , dephasing does not occur; when  $\eta = 0$ , there has been complete dephasing and  $\hat{\rho} = a|0\rangle\langle 0| + (1-a)|1\rangle\langle 1|$  for some  $0 \leq a \leq 1$ .

It is known that the optimal state for estimation of phase and the optimal state for estimation of dephasing both correspond to any equatorial state. We will select  $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ . After a phase rotation and dephasing, this yields

$$\hat{\rho} = \begin{pmatrix} \frac{1}{2} & \frac{\eta}{2}e^{i\phi} \\ \frac{\eta}{2}e^{-i\phi} & \frac{1}{2} \end{pmatrix}. \quad (5.24)$$

The eigenvalues of this are  $\frac{1 \pm \eta}{2}$  and the corresponding eigenvectors are  $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm e^{-i\phi}|1\rangle)$ . Referring to section 5.3.1, it is thus immediate that we *can* always simultaneously estimate phase and dephasing. One parameter is ‘quasi-classical’ and the other unitary. However, since we have a mixed state, this may not be achievable with measurements on individual qubits. In fact, a quick test shows that the optimal separable measurements for phase and dephasing yield no information about the opposite parameter.

Remarkably, if we consider POVMs along the equator, then there is a complete trade-off between the optimal measurement for estimating phase and the optimal measurement for estimating dephasing [154]. This reveals the importance of collective measurement for achieving the Cramér-Rao bound for mixed states.

We have proved that if we consider only separable  $\hat{\rho}^{\otimes N}$ , it is possible to simultaneously, independently and optimally estimate phase and dephasing if collective measurements are permitted. We are interested in expanding the class of allowed states, however, to include entangled states. Huelga et al. [161] showed that in this case, the optimal estimation for phase in the presence of dephasing is achieved with some non-maximally entangled state.

It is desirable then to test if simultaneous estimability can be achieved in some subset of the entangled states. To check this we take inspiration from [161], in which they find the optimal states for estimation of phase in the presence of dephasing are highly symmetric, exhibiting both permutational symmetry of the qubits, and also a parity symmetry under qubit flips, i.e. they are invariant under  $\sigma_x^{\otimes N}$  where  $N$  refers to the number of qubits.

## 5. MULTIPARAMETER METROLOGY

---

Due to the high degree of symmetry, it is convenient to shift our consideration to the angular momentum notation. On  $N$  qubits, the maximum total angular momentum is  $J = \frac{N}{2}$ . In general, we write  $|j, m\rangle$  to denote a general angular momentum eigenstate where  $0 \leq j \leq \frac{N}{2}$  and  $j$  goes between these limits in integer steps and similarly  $-j \leq m \leq j$ , where  $m$  also increases in integer steps.

It is known that permutational symmetry on  $N$  qubits allows for convenient decomposition of the density matrix into blocks of constant  $j$  due to Schur-Weyl duality [163, 164]. For example, for two qubits, permutationally symmetric states  $\hat{\rho}$  have a block representation given by,

$$\begin{pmatrix} \bullet & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & 0 \\ \bullet & \bullet & \bullet & 0 \\ 0 & 0 & 0 & \bullet \end{pmatrix} \quad (5.25)$$

where the  $3 \times 3$  submatrix is spanned by the triplet states,  $|1, 1\rangle = |\uparrow\uparrow\rangle$ ,  $|1, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$  and  $|1, -1\rangle = |\downarrow\downarrow\rangle$ . The bottom right entry is the lone singlet, which is antisymmetric under exchange of spins  $|0, 0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ . To avoid confusion with the angular momentum notation, on the right side of the equation we have shifted from the usual notation of a qubit in the computational basis to up and down spins, via  $|\uparrow\rangle = |1\rangle$  and  $|\downarrow\rangle = |0\rangle$ .

A particularly useful construction for the decomposition for a dephased state is found in the appendix of [165], where an explicit form is provided by means of the Clebsch-Gordan coefficients. To reproduce this, consider the action of a completely independent dephasing channel on a completely permutationally symmetric pure state. Such an initial state is described by the wavefunction  $|\psi_0\rangle = \sum_{m=-J}^J d_m |J, m\rangle$  where  $\sum_m |d_m|^2 = 1$ . Then the action of the dephasing channel  $\Lambda$  is written as:

$$\Lambda(\hat{\rho}_0) = \sum_{k=0}^N \sum_{\pi_k^N} \pi_k^N [K_1^{\otimes k} \otimes K_2^{\otimes N-k}] \cdot \hat{\rho}_0 \cdot \pi_k^N [K_1^{\dagger \otimes k} \otimes K_2^{\dagger \otimes N-k}]. \quad (5.26)$$

In the above equation, the terms  $\pi_k^N [K_1^{\otimes k} \otimes K_2^{\otimes N-k}]$  refer to permutations of the  $k$  and  $N - k$  Kraus operators, and we sum over all possible such permutations.

The crux of the construction is that it is possible to rewrite  $|J, m\rangle$  using the Clebsch-Gordan decomposition [166]. Setting  $j_1 = \frac{k}{2}$  and  $j_2 = \frac{N-k}{2}$  and denoting the Clebsch-Gordan coefficients by the non-standard notation  $\langle \frac{k}{2}, m_1, \frac{N-k}{2}, m_2 | j, m \rangle = c_{m_1 m_2}^{j, k, N}$ ,

$$\begin{aligned} \sigma_z^{\otimes k} \otimes \mathbb{1}^{\otimes N-k} |J, m\rangle &= \sum_{m_1} (-1)^{\frac{k}{2}-m_1} c_{m_1 m-m_1}^{J, k, N} \left| \frac{k}{2}, m_1, \frac{N-k}{2}, m-m_1 \right\rangle \\ &= \sum_{m_1, j, \alpha_j} (-1)^{\frac{k}{2}-m_1} c_{m_1 m-m_1}^{J, k, N} c_{m_1 m-m_1}^{j, k, N} |j, m, \alpha_j\rangle, \end{aligned}$$

where we have achieved the second step by inserting the resolution of the identity  $\sum_{m=-j}^j \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{\alpha_j} |j, m, \alpha_j\rangle \langle j, m, \alpha_j|$  and then utilising the orthogonality relations to combine the angular momenta appropriately. Here  $\alpha_j$  indexes the multiplicity of the  $j$  subspace.

Substituting  $\sigma_z$  and  $\mathbb{1}$  with the Kraus operators  $K_1$  and  $K_2$  respectively, and performing the above step for all  $|J, m\rangle \langle J, m'|$ , the resultant form of the dephased  $\hat{\rho} \equiv \hat{\rho}_\eta$  is

$$\begin{aligned} \hat{\rho} &= \sum_{k=0}^N \sum_{j, j'=|\frac{N}{2}-k|}^{\frac{N}{2}} \sum_{m, m'} \sum_{\alpha_j, \alpha_{j'}} d_m d_{m'}^* \left( \frac{1-\eta}{2} \right)^k \left( \frac{1+\eta}{2} \right)^{N-k} \times \\ &\quad C_{j, m}^{N, k} C_{j', m'}^{N, k} \sum_{\pi_k^N} \pi_k^N (|j, m, \alpha_j\rangle \langle j', m', \alpha_{j'}|) \\ &= \sum_{k=0}^N \sum_{j=|\frac{N}{2}-k|}^{\frac{N}{2}} \sum_{m, m'} \sum_{\alpha_j, \alpha_{j'}} \binom{N}{k} d_m d_{m'}^* \left( \frac{1-\eta}{2} \right)^k \left( \frac{1+\eta}{2} \right)^{N-k} \times \\ &\quad C_{j, m}^{N, k} C_{j', m'}^{N, k} |j, m\rangle \langle j, m'|, \end{aligned}$$

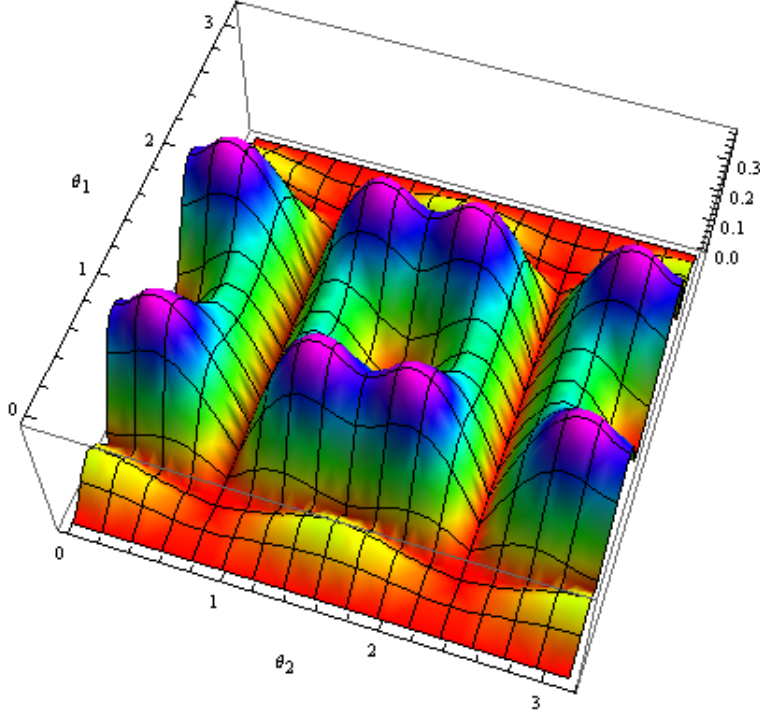
We have defined new coefficients  $C_{j, m}^{N, k} = \sum_{m_1=-\frac{k}{2}}^{\frac{k}{2}} (-1)^{\frac{k}{2}-m_1} c_{m_1 m-m_1}^{J, k, N} c_{m_1 m-m_1}^{j, k, N}$ . In the last step we have taken advantage of the permutation symmetry and Schur-Weyl duality. Furthermore, since all the multiplicity subspaces behave identically under our operations, we only need to look at one representative, thus eliminating the index  $\alpha_j$ . Such a simplification dramatically reduces computation time. The last step is to implement the phase unitary. This is easy, since it depends purely on  $m$ , so we need only substitute  $|j, m\rangle$  with  $e^{-i\frac{\phi}{2}m} |j, m\rangle$  in the above equation.

We use this construction to perform numerical simulations of dephasings. Unfortunately, it becomes quite quickly clear that permutationally symmetric states do *not* in

## 5. MULTIPARAMETER METROLOGY

---

general saturate the SLD Cramér-Rao bound. The absolute value of Equation (5.19) is plotted in fig 5.7 for the simplest case of two qubits with permutational symmetry.



**Figure 5.7:** A plot of Equation (5.19) for  $\eta = \frac{1}{2}$  and any value of  $\phi$  for the two-qubit state with  $d_0 = \cos \theta_1$ ,  $d_1 = \sin \theta_1 \cos \theta_2$  and  $d_{-1} = \sin \theta_1 \sin \theta_2$

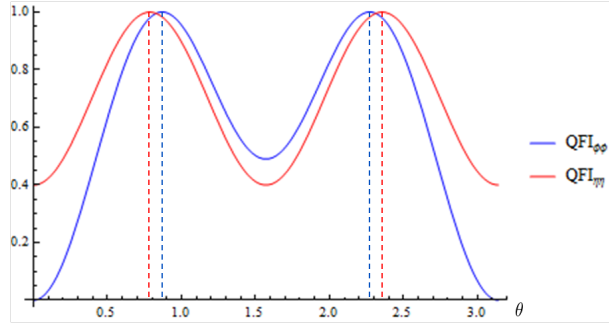
However, not all hope is lost, as can be seen by the 0s in the plot. As already mentioned, there is an additional symmetry which occurs in the optimal state for phase estimation in the presence of dephasing that we may be able to take advantage of. We will show that when the state is invariant under  $\sigma_x^{\otimes N}$ , we recover simultaneous and independent estimability. The two large grooves in Figure 5.7 at  $\theta_2 = \frac{\pi}{4}, \frac{3\pi}{4}$  represent such states (notably though, these are not the only states which allow simultaneous and independent estimability).

To prove this, we first calculate a further simplification on the structure of  $\hat{\rho}_\eta$ , *before* implementing the phase evolution (this is permitted because the actions of phase and dephasing commute). The parity symmetry implies that within each of the blocks of constant  $j$ , there exists a further splitting according to the irreducible representations of the parity operator.

The parity operator only has one-dimensional irreducible representations corresponding to the trivial and to the alternating representation. The eigenvectors of  $\hat{\rho}_\eta$  can then be chosen to have either even or odd parity. Given the block diagonal structure, the  $i^{\text{th}}$  even parity vector in the  $j$  subspace can be expressed as  $|\psi'_{\text{even},i}\rangle = \sum_m e_{i,m}^j (|j, m\rangle + |j, -m\rangle)$ , where  $\sum_m |e_{i,m}^j|^2 = 1$ . Similarly, all odd parity eigenvectors will have the structure  $|\psi'_{\text{odd},i}\rangle = \sum_m o_{i,m}^j (|j, m\rangle - |j, -m\rangle)$ , where  $\sum_m |o_{i,m}^j|^2 = 1$ .

Now consider the decomposition of the density matrix in terms of such eigenvectors  $\hat{\rho}_\eta = \sum_i p_i |\psi'_i\rangle \langle \psi'_i|$ . The unitary phase only serves to alter the eigenstates. Thus, after the phase unitary the density matrix is  $\hat{\rho}_{\eta,\phi} = \sum_i p_i |\psi_i\rangle \langle \psi_i|$ , where  $|\psi_i\rangle = U(\theta) |\psi'_i\rangle$ . Due to this  $\langle \psi_i | \partial_\eta \psi_k \rangle = \langle \psi_i | U^\dagger(\phi) \partial_\eta (U(\phi) |\psi_k\rangle) = \langle \psi'_i | \partial_\eta \psi'_k \rangle$ . This simplifies calculation of  $\langle \psi_i | \partial_\eta \hat{\rho}_{\eta,\phi} | \psi_k \rangle$  terms when  $i \neq k$ . Most significantly, one can observe that for  $\psi'_i$  and  $\psi'_k$  from subspaces corresponding to different parities,  $\langle \psi_i | \partial_\eta \psi_k \rangle = 0$ .

This is because the subspaces as a whole do not change with  $\eta$ . Returning to the almost-trivial example of the decomposition of two qubits into triplets and singlets, the singlet space always remains completely separate from the triplet space and it will not overlap with any combination of triplets regardless of  $\eta$ .



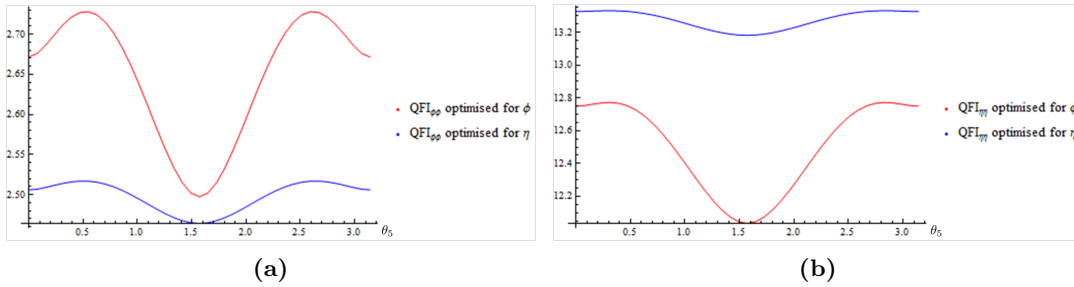
**Figure 5.8: Incompatibility of optimal states.** Plots of the fisher information (normalised to a maximum of 1) of phase and dephasing for  $N = 2$ . Recalling the form of the input states  $|\psi_0\rangle = \sum_m d_m |J, m\rangle$ , we set  $d_{-1} = d_1 = \frac{1}{\sqrt{2}} \cos(\theta)$  and  $d_0 = \sin(\theta)$ , and vary  $\theta$  along the  $x$ -axis. Clearly the maxima as indicated by the dashed lines do not agree.

Having eliminated approximately half of the terms of  $\text{Tr} \hat{\rho}_{\eta,\phi} L_\phi L_\eta$ , we turn our focus to the terms which include  $|\psi_i\rangle$  and  $|\psi_j\rangle$  from the *same* parity subspace. We will treat the even parity case, but the proof for odd parity is identical. After the phase rotation, the eigenstates become  $|\psi_{\text{even},i}\rangle = \sum_m e_{i,m}^j (e^{-i\frac{\phi}{2}m} |j, m\rangle + e^{i\frac{\phi}{2}m} |j, -m\rangle)$ . Differentiating this state with respect to  $\phi$  induces a sign difference between the two terms sharing

## 5. MULTIPARAMETER METROLOGY

the coefficient  $e_{i,m}^j$ . Using the orthonormality of  $|j, m\rangle$  gives us  $\langle\psi_{\text{even},i}|\partial_\phi\psi_{\text{even},k}\rangle = \sum_m e_{i,m}^{j,*} e_{k,m}^j (m - m) = 0$ . Thus every numerator term,  $\langle\psi_i|\partial_\eta\hat{\rho}|\psi_j\rangle\langle\psi_j|\partial_\phi\hat{\rho}|\psi_i\rangle$ , of  $\text{Tr}\hat{\rho}_{\eta,\phi}L_\phi L_\eta$  is equal to 0 and we can simultaneously estimate the parameters.

We can use the result presented earlier to explore numerically the optimal states for estimation of  $\phi$  and  $\eta$ . In Figure 5.8, the disagreement in optimal state is plotted. Although they are fairly close, it is clear that they do not match. Larger dimensional calculations still contain such a discrepancy. In Figure 5.9, we plot graphs for 10 qubits. We compose the constants  $d_i$  using products of trigonometric functions, such that  $\sum_i |d_i|^2 = 1$  as necessary. We then optimise all the parameters but one, and vary this from 0 to  $\pi$  for the plots. It is apparent that although the disparity in optimal state remains, it is not extremely harmful. Conveniently, the optimal state for estimating dephasing is always a separable state (although we do not have an analytical proof for this, it is indicated by numerics). If we choose a fully separable state then in the case of 10 qubits, we still achieve a QFI of  $> 90\%$  the optimal one for phase.



**Figure 5.9: Optimised plots for phase and dephasing.** In (a) we plot the quantum Fisher information for phase when all but one parameter  $d_i$  is optimised and we vary this final parameter along the x-axis. For the red line, these parameters are optimised for phase, with the blue line they are optimised for dephasing. Noting that the  $y$ -axis begins at about a value of 2.45, it is apparent that there is not a huge difference between the optimal states for estimating phase and dephasing. (b) We now plot the values for dephasing in the same manner as for phase: for the blue line all but one parameter is optimised for dephasing and for the red, all but one is optimised for phase. We see similar results in that there is not a significant difference between optimal values.

The optimal state for joint estimation thus depends on the weighting of importance between dephasing and phase estimation. For practical purposes, however, it seems likely that independent estimation would be optimal simply for comparative ease of

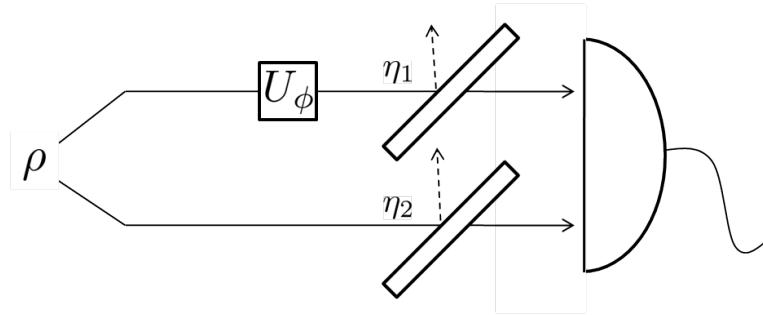


implementation, especially considering the caveats of section 5.1.2. This would still require large, collective measurements, however, meaning that it could not be done as a scheme involving repeated measurements on single qubits.

To summarise, in the case of phase estimation in the presence of independent dephasing, we have identified that states which are both permutationally and spin-flip symmetric are always suitable probes for simultaneously and independently estimating both parameters. The optimal states for estimating each parameter are not identical, although optimising the input state for one parameter is not hugely detrimental to the estimation of the other.

#### 5.3.4 Interferometry with Symmetric Loss

An interesting thing to search for is a lossy model in which the SLDs properly commute (not just in expectation value). This model was stumbled upon largely by luck, after altering a scheme found in [167]. In that paper they attempted to estimate both phase and loss in a single arm interferometer but found that the commutation condition could not be satisfied, even though the quantum Fisher information matrix was always diagonal.



**Figure 5.10: An interferometer with loss in both arms and a phase shift  $\phi$  in one of the arms.** We study the case of symmetric loss where the transmissivities are given by  $\eta_1 = \eta_2 = \eta$ .

To derive our example, we model the lossy interferometer of Figure 5.10 with an input pure state. It should be noted that this sort of model is not likely to be extremely well motivated given that any object which induces a phase shift is likely to bring with

## 5. MULTIPARAMETER METROLOGY

---

it additional loss. An input state of same constant photon number can be written as

$$|\psi\rangle = \sum_{k=0}^N \alpha_k |k, N-k\rangle. \quad (5.27)$$

After passing through the interferometer the resultant state is

$$|\psi\rangle = \sum_{k=0}^N \sum_{l_2=0}^{N-k} \sum_{l_1=0}^k \alpha_k e^{ik\phi} |k-l_1, N-k-l_2\rangle \otimes |l_1, l_2\rangle, \quad (5.28)$$

where  $l_{1,2}$  refers to the number of photons lost from the first and second branch of the scheme respectively. On tracing out the auxiliary modes, we obtain a density matrix

$$\hat{\rho} = \sum_{l_2=0}^{N-l_1} \sum_{l_1=0}^N p_{l_1 l_2} |\psi_{l_1 l_2}\rangle \langle \psi_{l_1 l_2}|, \quad (5.29)$$

where

$$|\psi_{l_1 l_2}\rangle = \frac{1}{\sqrt{p_{l_1 l_2}}} \sum_{k=l_1}^{N-l_2} \alpha_k e^{ik\phi} \sqrt{B_{l_1 l_2}^k} |k-l_1, N-k-l_2\rangle, \quad (5.30)$$

and

$$\begin{aligned} B_{l_1 l_2}^k &= \binom{k}{l_1} \binom{N-k}{l_2} \eta_1^{k-l_1} (1-\eta_1)^{l_1} \eta_2^{N-k-l_2} (1-\eta_2)^{l_2} \\ &= \binom{k}{l_1} \binom{N-k}{l_2} \eta^{N-l_1-l_2} (1-\eta)^{l_1+l_2} \quad \text{for } \eta_1 = \eta_2, \end{aligned} \quad (5.31)$$

are coefficients found by considering a binomial distribution for the loss of photons where the probability of successful passage, i.e. the transmissivity, is denoted by  $\eta$ . Lastly

$$p_{l_1 l_2} = \sum_{k=l_1}^{N-l_2} |\alpha_k|^2 B_{l_1 l_2}^k. \quad (5.32)$$

We are primarily interested at the moment in calculating the loss SLD. It is apparent that for fixed total loss the eigenstates  $\langle \psi_{l_1, L-l_1} | \psi_{l'_1, L'-l'_1} \rangle = 0$  for  $L \neq L'$ , but not necessarily for  $L = L'$ . This is clear because different total loss implies that we are left

with different total photon number states which are necessarily orthogonal. We can thus rewrite the loss matrix with the parametrization:

$$\hat{\rho} = \bigoplus_L \sum_{l_1} p_{l_1, L-l_1} |\psi_{l_1, L-l_1}\rangle \langle \psi_{l_1, L-l_1}|$$

If we absorb the probabilities into the wavefunctions (so that they're unnormalised), it's clear that the only dependence on  $\eta$  is in the terms  $B_{l_1, L-l_1}^k$ . An important calculation for this is, fixing  $l_1 + l_2 = L$ ,

$$\begin{aligned} \frac{d}{d\eta} B_{l_1, L-l_1}^k &= \binom{k}{l_1} \binom{N-k}{L-l_1} ((N-L)\eta^{N-L-1}(1-\eta)^L - L\eta^{N-L}(1-\eta)^{L-1}), \\ &= \binom{k}{l_1} \binom{N-k}{L-l_1} \left( \frac{N-L}{\eta} - \frac{L}{1-\eta} \right) \eta^{N-L}(1-\eta)^L, \\ &= \left( \frac{N-L}{\eta} - \frac{L}{1-\eta} \right) B_{l_1, L-l_1}^k, \\ &\equiv C_{NL}^\eta B_{l_1, L-l_1}^k. \end{aligned} \tag{5.33}$$

thus finally

$$\frac{\partial}{\partial \eta} \hat{\rho} = \bigoplus_L C_{NL}^\eta \sum_{l_1} p_{l_1, L-l_1} |\psi_{l_1, L-l_1}\rangle \langle \psi_{l_1, L-l_1}|.$$

It is useful that the block diagonal structure of the density matrix is maintained since the derivative with respect to phase will not mix states of different total photon number,  $\langle \psi_{l'_1, L'-l'_1} | \partial_\phi \psi_{l_1, L-l_1} \rangle = 0$  when  $L \neq L'$ . Thus, in searching for a mutual eigenbasis for the optimal measurement, we need only optimise within each block of the density matrix.

Continuing to focus especially on the loss, a brief calculation show that owing to the block structure the SLD can be written as  $L_\eta = \bigoplus_L \sum_{l_1} |\psi_{l_1, L-l_1}\rangle \langle \psi_{l_1, L-l_1}|$ . In each block of constant  $L$ ,  $\sum_{l_1} |\psi_{l_1, L-l_1}\rangle \langle \psi_{l_1, L-l_1}|$  is the identity and thus it appears that the optimal measurement for loss is simply projecting onto the identity for each block of constant  $L$ . To be even more explicit, we can write this POVM as  $\{\Pi_L\}$  with  $\Pi_L = \sum_{k=0}^{N-L} |k, N-k-L\rangle \langle k, N-k-L|$ ; since we are interested in knowing the number of photons lost, we project onto blocks of constant *remaining* photons. The resultant Fisher information will be,  $\sum_L C_{NL}^2 P(L|\eta)$  where  $P(L|\eta) = \text{Tr } \hat{\rho} \Pi_L = \sum_{l_1} p_{l_1, L-l_1}$ .

## 5. MULTIPARAMETER METROLOGY

---

Importantly, it can be shown that  $P(L|\eta) = \eta^{N-L}(1-\eta)^L \binom{N}{L}$ , which is state independent. Intuitively, this is the result we expect since distinguishing which arm photons were lost from is not beneficial given the symmetry of the lossiness. We therefore simply get a binomial distribution of photon loss, and sampling this is the most informative thing we can do.

Because  $\partial_\phi \hat{\rho}$  maintains the block diagonal structure,  $L_\phi$  can also be written in the same diagonal blocks as  $\hat{\rho}$  and  $L_\eta$ . Given that the elements of the optimal measurement for estimating loss are projections onto blocks of constant photon number, this means that, it acts as the identity on each block and necessarily commutes with  $L_\phi$ .

The exact measurement strategy for phase estimation is difficult to derive apart from in very specific cases. This, as well as the bounds on variance of the phase estimation, are problems which have been addressed in [168]. The main advantage here is that we discover the loss estimation is obtained completely for free, with no regard to the input state (as long as it is of constant total photon number) and we can thus use known strategies for estimating the phase.

As a final note, we point out an interesting conceptual point: for an interferometer with loss and phase in only a single arm, so that  $\eta_2 = 1$  in Figure 5.10, the commutation condition is satisfied only when all photons are passed through one arm [167]. This means that only if we gain no information at all about the phase can we satisfy the SLD Cramér-Rao bound. However, if we were to add an equivalent loss to the second arm, then we gain exactly the same information about the loss parameter, while also being able to optimise the estimation for maximal information on the phase. While on first thought this might seem a little bit surprising that adding loss to one arm can cause such a dramatic improvement, on second thought, it becomes obvious why this should be the case. In losing photons, we *gain* information about the loss parameter; we learn from the lack of photons, as well as their presence.

### 5.3.5 Summary

This chapter of the thesis has focussed on clarifying the concepts of multiparameter estimation, which though mathematically precise have been occasionally blurry in the physics literature. We have also provided illustrative examples to various scenarios one might encounter.

The emphasis has largely been on how we can reduce problems to the familiarity of the SLD formalism of single parameter estimation. This renders the derived bounds asymptotically achievable (although possibly with difficulty), and thus quite informative. To do this, we need to choose states which allow us to side-step the problems which can arise due to the non-commutativity of measurements in the quantum case, which in the formalism we've outlined, corresponds to satisfying the SLD commutation condition. However, we choose to go even further and also require a diagonal Fisher information matrix, which is really mostly a classical consideration. For this formalism to be of any use we require exquisite control of our input state (and also measurement).

In the examples, this simplification has been especially successful in the case of two-parameter pure state estimation as well as the symmetrised lossy interferometer, where the measurements have been identified. In the case of phase and dephasing, the need for large collective measurements may be necessary which is a complicating factor, but calculating the ultimate bound for this open problem was still a useful result, given other work on the topic left this question open [154].

Detailed, realistic analyses of multiparameter estimation problems need to be performed. It would be an interesting experiment to demonstrate sub-shot noise scaling in two parameters at once, and perhaps the two wave-plate setup with entangled photon pairs would be feasible for achieving this.

## 6

# Conclusions

This brings the thesis to its end. We've covered a broad set of topics, loosely united by their practical functionality as useful for imaging and metrology, as well as their suitability to optical setups.

In the first non-introductory chapter, we considered how different concepts of quantumness could produce drastically different results on the role of quantumness in a system. We studied three intensity interferometric schemes, that of Hanbury Brown and Twiss, which was one of the earliest examples of intensity interferometry; that of ghost imaging, which reawakened a controversy on whether intensity interferometry could afford an interpretation in terms of classical intensity correlations; and lastly, that of quantum illumination, which, while not especially controversial in of itself, provided us with grounds to experimentally test some of the ideas in our analysis of intensity interferometry schemes.

The light used in all these schemes is described by a positive  $P$ -representation and can thus be regarded as classical from the usual definitions of quantum optics. Counter-intuitively this classicality holds regardless of the illumination conditions so that even in the photon-counting regime, the system can be regarded as classical. However, defining quantumness using the quantum discord reveals a completely different result, which is that low illumination intensity interferometry is indeed quite quantum, since the majority of the total correlations is made up by discord. This is consistent with previous observations that discord and the  $P$ -representation identify highly disparate sets of classical states and moreover, it's consistent with explicit physical models which

---

show that low-illumination easily lends itself to a picture involving some sort of non-local interference.

In the second part of the main body of the thesis, we focussed on a very different set of problems of estimating multiple parameters at the same time. This included estimation of two unitaries simultaneously as well as estimation of unitary parameters with decoherence parameters. We clarified some uncertainties present in the physics literature on parameter estimation and laid out a framework of criteria which, when satisfied, render multiparameter problems effectively analogous to the more familiar case of single-parameter estimation (especially for pure states).

Amongst the interesting results gleaned is that – to an extent – Heisenberg scaling can be achieved in two parameters simultaneously. Further, phase and dephasing can both be estimated simultaneously, although their optimal states differ. As far as practical considerations go, it appears that separable states will always be optimal for this case. Lastly, we also found that phase and symmetric loss can always be estimated simultaneously with simple measurements in a Mach-Zehnder interferometer.

From both the sections of the thesis we are left with open questions and potential research lines. In the first part, alternative models of coarse-graining are possible. A very promising avenue for this is applicable in the relevant regime of large mode number per pixel where the central limit theorem leads to Gaussian photon-counting distributions on the detectors, which can be approximated by ‘thermalised’ coherent states of appropriate brightness (by noting that the Poissonian distribution with high mean approximates a Gaussian, with the thermalisation allowing for different variances). It seems feasible that this would demonstrate that effective classicality naturally arises from the collective counting of photons. Moreover, it would be interesting to use alternative formulations of quantum correlations which more accurately correspond to the measurements performed, such as the measurement induced disturbance [169] which considers discord-like quantum correlations with photon-counting measurements, rather than optimisation over all measurements.

In the metrology section, adapting the schemes for realistic considerations would be a sensible next step. In particular, the two-unitary scheme was assumed to be decoherence-free. It may be worthwhile to realistically model this scheme to see if a sub-shot-noise scaling in two parameters is feasible in an experimental setting. Moreover, some of the most significant applications of the results of quantum estimation theory use

## 6. CONCLUSIONS

---

continuous variables as tools, such as the LIGO gravitational wave detectors [170]. We have only considered qubits here, and most multiple parameter estimation problems in the wider literature only consider qubit inputs. Extending the consideration of multiple parameter estimation schemes to continuous variable inputs could open the door to practical implementations.



## Appendix A

# Gaussian Moment Factoring

Gaussian moment factoring, alternatively known as Isserlis' theorem, is a fundamental result which arises from the fact that all moments of a Gaussian probability distribution are expressible in terms of its second and first moments. To clarify this, for a set of  $m$  classical Gaussian variates  $\{x_1, \dots, x_m\}$ , we can define a general central moment of these by

$$\langle \Delta x_{i_1} \dots \Delta x_{i_n} \rangle = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \sum_{\text{all pairings}} \langle \Delta x_{i_1} \Delta x_{i_2} \rangle \dots \langle \Delta x_{i_{n-1}} \Delta x_{i_n} \rangle & \text{if } n \text{ is even} \end{cases} \quad (\text{A.1})$$

The indices  $i_k \in \{1, 2, \dots, m\}$  may be repeated.

Following [68], this can be proven by first taking a set of independent variates  $\{y_{i_k}\}$ , such that any moment  $\langle \Delta y_{i_1} \Delta y_{i_2} \dots \Delta y_{i_n} \rangle$  is only non-zero if every index appears an even number of times. If this is indeed the case, we can write

$$\langle (\Delta y_{i_1})^{2l_{i_1}} (\Delta y_{i_2})^{2l_{i_2}} \dots (\Delta y_{i_n})^{2l_{i_n}} \rangle = \quad (\text{A.2})$$

where  $l_{i_k}$  are positive integers and  $n!! = (n)(n-2)(n-4)\dots$ . This can be easily deduced from the moment generating function for single (or independent) variates, from which we find  $\langle \Delta y^{2l} \rangle = (2l-1)!! \langle (\Delta y)^2 \rangle$ .

From the case of independent variates  $\{y_i\}$  we can deduce Equation (A.1) by noting that we can relate these to arbitrary Gaussian variates by a unitary  $U$  transformation, so that  $x_i = \sum_j U_{ij} y_j$ .

## A. GAUSSIAN MOMENT FACTORING

---

It is non-trivial to extend this theorem to the case of our non-commuting mode operators, however, this can be done by using anti-normally ordered operators and the theory of heterodyne detection [171, 172], along with the classical moment factoring theorem. Subsequently, we may rearrange the operators into any ordering by using the commutation relations and we find the moment factoring theorem operates much as in the classical case. For example, the second-order coherence function can be decomposed as

$$\langle \hat{a}_1^\dagger \hat{a}_1 \hat{a}_2^\dagger \hat{a}_2 \rangle = \langle \hat{a}_1^\dagger \hat{a}_1 \rangle \langle \hat{a}_2^\dagger \hat{a}_2 \rangle + \langle \hat{a}_1^\dagger \hat{a}_2 \rangle \langle \hat{a}_2^\dagger \hat{a}_1 \rangle + \langle \hat{a}_1^\dagger \hat{a}_2^\dagger \rangle \langle \hat{a}_1 \hat{a}_2 \rangle \quad (\text{A.3})$$

For phase-insensitive light (such as thermal states), the terms  $\langle \hat{a}_1^\dagger \hat{a}_2^\dagger \rangle$  and  $\langle \hat{a}_1 \hat{a}_2 \rangle$  are equal to 0. On the other hand, for light for which all the correlations are phase-dependent (such as two mode squeezed states), we find terms like  $\langle \hat{a}_2^\dagger \hat{a}_1 \rangle$  and  $\langle \hat{a}_1^\dagger \hat{a}_2 \rangle$  will be 0.

# References

- [1] Ragy, S. & Adesso, G. *Sci. Rep.*, **2** 651 (2012) (Cited on pages 1, 86, and 90.)
- [2] Ragy, S. & Adesso, G. *Physica Scripta*, **2013** T153 014052 (2013) (Cited on pages 1 and 93.)
- [3] Ragy, S., Berchera, I. R., Degiovanni, I. P., Olivares, S., Paris, M. G., Adesso, G. & Genovese, M. *JOSA B*, **31** 9 2045 (2014) (Cited on pages 1 and 87.)
- [4] Adesso, G., Ragy, S. & Lee, A. R. *Open Systems & Information Dynamics*, **21** 01n02 (2014) (Cited on pages 1, 2, 34, 45, and 46.)
- [5] Adesso, G., Ragy, S. & Girolami, D. *Classical and Quantum Gravity*, **29** 22 224002 (2012) (Cited on page 1.)
- [6] Kogias, I., Ragy, S. & Adesso, G. *Physical Review A*, **89** 5 052324 (2014) (Cited on pages 1 and 12.)
- [7] Kogias, I., Lee, A. R., Ragy, S. & Adesso, G. *Phys. Rev. Lett.*, **114** 060403 (2015) (Cited on page 1.)
- [8] Ferraro, A., Aolita, L., Cavalcanti, D., Cucchietti, F. & Acin, A. *Physical Review A*, **81** 5 052318 (2010) (Cited on pages 2 and 32.)
- [9] Lopaeva, E., Berchera, I. R., Degiovanni, I., Olivares, S., Brida, G. & Genovese, M. *Physical review letters*, **110** 15 153603 (2013) (Cited on pages 5, 87, 88, 89, 90, 91, 92, 93, 94, 95, and 96.)
- [10] Schrödinger, E. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 31, 555–563 (Cambridge Univ Press) (1935) (Cited on pages 10 and 19.)

## REFERENCES

---

- [11] McKeever, J., Boca, A., Boozer, A., Miller, R., Buck, J., Kuzmich, A. & Kimble, H. *Science*, **303** 5666 1992 (2004) (Cited on page 10.)
- [12] Nägerl, H. C., Leibfried, D., Rohde, H., Thalhammer, G., Eschner, J., Schmidt-Kaler, F. & Blatt, R. *Physical Review A*, **60** 1 145 (1999) (Cited on page 10.)
- [13] Wrachtrup, J. & Jelezko, F. *Journal of Physics: Condensed Matter*, **18** 21 S807 (2006) (Cited on page 10.)
- [14] Dowling, J. P. & Milburn, G. J. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **361** 1809 1655 (2003) (Cited on page 11.)
- [15] Shor, P. W. *SIAM journal on computing*, **26** 5 1484 (1997) (Cited on page 11.)
- [16] Giovannetti, V., Lloyd, S. & Maccone, L. *Nature Photonics*, **5** 4 222 (2011) (Cited on pages 11 and 105.)
- [17] Bennett, C. H. & Brassard, G. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, volume 175, 8 (1984) (Cited on page 11.)
- [18] Sudarshan, E. C. G. *Phys. Rev. Lett.*, **10** 277 (1963) (Cited on pages 11, 39, 42, and 60.)
- [19] Vidal, G. *Journal of Modern Optics*, **47** 2-3 355 (2000) (Cited on page 11.)
- [20] de Vicente, J. I. *Journal of Physics A: Mathematical and Theoretical*, **47** 42 424017 (2014) (Cited on page 11.)
- [21] Baumgratz, T., Cramer, M. & Plenio, B., M. *Phys. Rev. Lett.*, **113** 140401 (2014) (Cited on page 11.)
- [22] Brandão, F. G., Horodecki, M., Oppenheim, J., Renes, J. M. & Spekkens, R. W. *Physical review letters*, **111** 25 250404 (2013) (Cited on page 11.)
- [23] Jozsa, R. & Linden, N. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **459** 2036 2011 (2003) (Cited on pages 11 and 19.)

## REFERENCES

---

- [24] Datta, A., Shaji, A. & Caves, C. M. *Physical review letters*, **100** 5 050502 (2008)  
(Cited on page 11.)
- [25] Girolami, D., Souza, A. M., Giovannetti, V., Tufarelli, T., Filgueiras, J. G., Sarthour, R. S., Soares-Pinto, D. O., Oliveira, I. S. & Adesso, G. *Physical Review Letters*, **112** 21 210401 (2014) (Cited on page 11.)
- [26] Shapiro, J. H. & Boyd, R. W. *Quantum Information Processing*, **11** 4 949 (2012)  
(Cited on pages 11, 60, 68, 71, 72, and 85.)
- [27] Hanbury Brown, R. & Twiss, R. Q. *Nature* (Cited on pages 12, 59, and 61.)
- [28] Helstrom, C. W. *Journal of Statistical Physics*, **1** 2 231 (1969) (Cited on pages 12, 99, and 103.)
- [29] Holevo, A. S. *Probabilistic and statistical aspects of quantum theory* (North-Holland) (1982) (Cited on pages 12, 99, 102, 103, and 109.)
- [30] Boixo, S., Datta, A., Davis, M. J., Flammia, S. T., Shaji, A. & Caves, C. M. *Phys. Rev. Lett.*, **101** 040403 (2008) (Cited on page 12.)
- [31] Zwierz, M., Pérez-Delgado, C. A. & Kok, P. *Phys. Rev. Lett.*, **105** 180402 (2010)  
(Cited on page 12.)
- [32] Gendra, B., Ronco-Bonvehi, E., Calsamiglia, J., Muñoz-Tapia, R. & Bagan, E. *Physical review letters*, **110** 10 100501 (2013) (Cited on page 12.)
- [33] Gendra, B., Calsamiglia, J., Muñoz-Tapia, R., Bagan, E. & Chiribella, G. *Physical Review Letters*, **113** 26 260402 (2014) (Cited on page 12.)
- [34] Combes, J., Ferrie, C., Jiang, Z. & Caves, C. M. *Phys. Rev. A*, **89** 052117 (2014)  
(Cited on page 12.)
- [35] Pusey, M. F., Barrett, J. & Rudolph, T. *Nature Physics*, **8** 6 475 (2012) (Cited on page 13.)
- [36] Lewis, P. G., Jennings, D., Barrett, J. & Rudolph, T. *Physical review letters*, **109** 15 150404 (2012) (Cited on page 13.)

## REFERENCES

---

- [37] Colbeck, R. & Renner, R. *arXiv preprint arXiv:1312.7353* (2013) (Cited on page 13.)
- [38] Nielsen, M. A. & Chuang, I. L. *Quantum computation and quantum information* (Cambridge university press) (2010) (Cited on pages 13, 17, 21, and 33.)
- [39] Wilde, M. M. *Quantum information theory* (Cambridge University Press) (2013) (Cited on pages 13 and 33.)
- [40] Winter, A. *Journal of Physics A: Mathematical and General*, **34** 35 7095 (2001) (Cited on page 14.)
- [41] Breuer, H. & Petruccione, F. *The Theory of Open Quantum Systems* (OUP Oxford) (2007) (Cited on page 17.)
- [42] Wiseman, H. M. *Journal of Physics A: Mathematical and Theoretical*, **47** 42 424001 (2014) (Cited on page 20.)
- [43] Maudlin, T. *Journal of Physics A: Mathematical and Theoretical*, **47** 42 424010 (2014) (Cited on page 20.)
- [44] Werner, R. F. *Journal of Physics A: Mathematical and Theoretical*, **47** 42 424011 (2014) (Cited on page 20.)
- [45] Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. *Phys. Rev. Lett.*, **23** 880 (1969) (Cited on page 20.)
- [46] Preskill, J. Lecture Notes for Ph219/CS219: Quantum Information and Computation Chapter 4 (Cited on page 20.)
- [47] Bohm, D. *Phys. Rev.*, **85** 166 (1952) (Cited on page 21.)
- [48] Bohm, D. *Phys. Rev.*, **85** 180 (1952) (Cited on page 21.)
- [49] Bohm, D., Hiley, B. & Kaloyerou, P. *An Ontological Basis for the Quantum Theory*. Physics reports (North-Holland) (1987) (Cited on page 21.)
- [50] Chiribella, G., D’Ariano, G. M. & Perinotti, P. *Phys. Rev. A*, **84** 012311 (2011) (Cited on page 21.)

- [51] Escher, B. M., de Matos Filho, R. L. & Davidovich, L. *Nature Physics*, **7** 5 406 (2011) (Cited on page 22.)
- [52] Shannon, C. E. *ACM SIGMOBILE Mobile Computing and Communications Review*, **5** 1 3 (2001) (Cited on page 26.)
- [53] Plenio, M. B. & Virmani, S. *arXiv preprint quant-ph/0504163* (2005) (Cited on page 28.)
- [54] Ollivier, H. & Zurek, W. H. *Physical review letters*, **88** 1 017901 (2001) (Cited on pages 30, 31, and 85.)
- [55] Henderson, L. & Vedral, V. *Journal of Physics A: Mathematical and General*, **34** 35 6899 (2001) (Cited on pages 30, 31, and 85.)
- [56] Werner, R. F. *Physical Review A*, **40** 8 4277 (1989) (Cited on page 30.)
- [57] Streltsov, A., Kampermann, H. & Bruß, D. *Physical review letters*, **107** 17 170502 (2011) (Cited on page 31.)
- [58] Modi, K., Brodutch, A., Cable, H., Paterek, T. & Vedral, V. *Rev. Mod. Phys.*, **84** 1655 (2012) (Cited on pages 31 and 85.)
- [59] Rényi, A. *Probability theory*. North-Holland series in applied mathematics and mechanics (Elsevier) (1970) (Cited on page 32.)
- [60] Adesso, G., Girolami, D. & Serafini, A. *Physical review letters*, **109** 19 190502 (2012) (Cited on pages 33 and 54.)
- [61] Ketterer, A., Walborn, S., Keller, A., Coudreau, T. & Milman, P. *arXiv preprint arXiv:1406.6388* (2014) (Cited on page 34.)
- [62] Lloyd, S. & Braunstein, S. L. In *Quantum Information with Continuous Variables*, 9–17 (Springer) (2003) (Cited on page 34.)
- [63] Simon, R., Sudarshan, E. C. G. & Mukunda, N. *Phys. Rev. A*, **36** 3868 (1987) (Cited on page 34.)
- [64] Adesso, G. & Illuminati, F. *Journal of Physics A: Mathematical and Theoretical*, **40** 28 7821 (2007) (Cited on pages 34 and 47.)

## REFERENCES

---

- [65] Schumaker, B. L. *Physics Reports*, **135** 6 317 (1986) (Cited on pages 34, 43, and 45.)
- [66] Weedbrook, C., Pirandola, S., Garcia-Patron, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H. & Lloyd, S. *Reviews of Modern Physics*, **84** 2 621 (2012) (Cited on page 34.)
- [67] Ferraro, A., Olivares, S. & Paris, M. G. *Gaussian states in quantum information* (Bibliopolis) (2005) (Cited on page 34.)
- [68] Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics* (Cambridge University Press) (1995) (Cited on pages 35, 42, 63, 64, 66, and 137.)
- [69] Loudon, R. *The Quantum Theory of Light* (OUP Oxford) (2000) (Cited on page 35.)
- [70] Cahill, K. E. & Glauber, R. J. *Phys. Rev.*, **177** 1857 (1969) (Cited on pages 36 and 41.)
- [71] Kolmogorov, A. *Foundations of the Theory of Probability*. (Cited on page 37.)
- [72] Tijms, H. *Understanding Probability* (Cambridge University Press) (2012) (Cited on page 37.)
- [73] Wigner, E. *Phys. Rev.*, **40** 749 (1932) (Cited on page 39.)
- [74] Moyal, J. E. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 45, 99–124 (Cambridge Univ Press) (1949) (Cited on page 39.)
- [75] Dirac, P. Letter to Moyal, 9 Jan 1946 (Cited on page 39.)
- [76] Husimi, K. *Nippon Sugaku-Buturigakkwai Kizi Dai 3 Ki*, **22** 4 264 (1940) (Cited on page 39.)
- [77] Glauber, R. *Phys. Rev. Lett.*, **10** 84 (1963) (Cited on pages 39 and 60.)
- [78] Leonhardt, U. *Measuring the Quantum State of Light*. Cambridge Studies in Modern Optics (Cambridge University Press) (2005) (Cited on pages 40 and 52.)
- [79] Bertrand, J. & Bertrand, P. *Foundations of Physics*, **17** 4 397 (1987) (Cited on page 40.)



- [80] Simon, R. *Phys. Rev. Lett.*, **84** 2726 (2000) (Cited on page 41.)
- [81] de Gosson, M. *Symplectic Geometry and Quantum Mechanics*. Operator Theory: Advances and Applications / Advances in Partial Differential Equations (Birkhäuser Basel) (2006) (Cited on pages 45 and 47.)
- [82] Lee, A. R. *arXiv preprint arXiv:1309.4419* (2013) (Cited on page 46.)
- [83] Williamson, J. *Am. J. Math.*, **58** 141 (1936) (Cited on page 46.)
- [84] Serafini, A. *Phys. Rev. Lett.*, **96** 110402 (2006) (Cited on page 47.)
- [85] Laurat, J., Keller, G., Oliveira-Huguenin, J. A., Fabre, C., Coudreau, T., Serafini, A., Adesso, G. & Illuminati, F. *Journal of Optics B: Quantum and Semiclassical Optics*, **7** 12 S577 (2005) (Cited on page 51.)
- [86] Gardiner, C. & Zoller, P. *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*. Springer Series in Synergetics (Springer) (2010) (Cited on page 52.)
- [87] Mišta, L., Tatham, R., Girolami, D., Korolkova, N. & Adesso, G. *Phys. Rev. A*, **83** 042325 (2011) (Cited on page 52.)
- [88] Eisert, J., Scheel, S. & Plenio, M. *Physical review letters*, **89** 13 137903 (2002) (Cited on page 53.)
- [89] Fiurášek, J. *Physical review letters*, **89** 13 137904 (2002) (Cited on page 53.)
- [90] Giedke, G. & Cirac, J. I. *Physical Review A*, **66** 3 032316 (2002) (Cited on page 53.)
- [91] Einstein, A., Podolsky, B. & Rosen, N. *Phys. Rev.*, **47** 777 (1935) (Cited on page 53.)
- [92] Horn, R. & Johnson, C. *Matrix Analysis*. Matrix Analysis (Cambridge University Press) (2012) (Cited on page 54.)
- [93] Pirandola, S., Spedalieri, G., Braunstein, S. L., Cerf, N. J. & Lloyd, S. *Physical review letters*, **113** 14 140405 (2014) (Cited on page 56.)

## REFERENCES

---

- [94] Adesso, G. & Datta, A. *Physical review letters*, **105** 3 030501 (2010) (Cited on pages 57 and 74.)
- [95] Hanbury Brown, R. & Twiss, R. Q. *Proc. R. Soc. Lond. A*, **242** 300 (1957) (Cited on pages 59 and 61.)
- [96] Brown, R. H. & Twiss, R. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, **243** 1234 291 (1958) (Cited on pages 59 and 61.)
- [97] Brannen, E. & Ferguson, H. *Nature*, **178** 481 (1956) (Cited on pages 59 and 61.)
- [98] Purcell, E. *Nature*, **178** 1449 (1956) (Cited on pages 59 and 61.)
- [99] Meyers, R. E., Deacon, K. S. & Shih, Y. *Applied Physics Letters*, **98** 11 111115 (2011) (Cited on page 60.)
- [100] Pittman, T., Shih, Y., Strekalov, D. & Sergienko, A. *Physical Review A*, **52** 5 R3429 (1995) (Cited on pages 60 and 69.)
- [101] Strekalov, D. V., Sergienko, A. V., Klyshko, D. N. & Shih, Y. H. *Phys. Rev. Lett.*, **74** 3600 (1995) (Cited on page 60.)
- [102] Kolobov, M. *Quantum Imaging* (Springer New York) (2007) (Cited on page 60.)
- [103] Lloyd, S. *Science*, **321** 5895 1463 (2008) (Cited on pages 60, 87, 89, 92, and 96.)
- [104] Tan, S.-H., Erkmen, B. I., Giovannetti, V., Guha, S., Lloyd, S., Maccone, L., Pirandola, S. & Shapiro, J. H. *Physical review letters*, **101** 25 253601 (2008) (Cited on pages 60, 87, 89, 90, 92, and 96.)
- [105] Shapiro, J. H. *Physical Review A*, **80** 2 022320 (2009) (Cited on page 60.)
- [106] Shih, Y. H. *Advances in Lasers and Electro Optics* (2010) (Cited on pages 60, 68, 78, and 79.)
- [107] Scarcelli, G., Berardi, V. & Shih, Y. H. *Phys. Rev. Lett.*, **96** 063602 (2007) (Cited on pages 60, 62, 68, 79, and 85.)
- [108] Gatti, A., Bondani, M., Lugiato, L. A., Paris, M. G. A. & Fabre, C. *Phys. Rev. Lett.*, **98** 039301 (2007) (Cited on pages 60 and 73.)

## REFERENCES

---

- [109] Hanbury Brown, R. & Twiss, R. Q. *Phil. Mag.*, **45** 663 (1954) (Cited on page 61.)
- [110] Gatti, A., Brambilla, E., Bache, M. & Lugiato, L. A. *Phys. Rev. Lett.*, **93** 093602 (2004) (Cited on pages 62, 70, and 90.)
- [111] Gabor, D. Lectures in Communication Theory. Technical Report 238, Research Lab of Electronics, M.I.T. (1952) (Cited on page 67.)
- [112] Fellgett, P. *Nature*, **179** 956 (Cited on page 67.)
- [113] Abouraddy, A. F., Saleh, B. E., Sergienko, A. V. & Teich, M. C. *Physical review letters*, **87** 12 123602 (2001) (Cited on pages 68, 69, and 70.)
- [114] Bennink, R. S., Bentley, S. J. & Boyd, R. W. *Physical review letters*, **89** 11 113601 (2002) (Cited on pages 68 and 70.)
- [115] Valencia, A., Scarcelli, G., D'Angelo, M. & Shih, Y. *Physical review letters*, **94** 6 063601 (2005) (Cited on page 68.)
- [116] Shih, Y. *Quantum Information Processing*, 1–7 (2012) (Cited on pages 68, 73, and 86.)
- [117] Ferraro, A. & Paris, M. G. *Physical review letters*, **108** 26 260403 (2012) (Cited on page 69.)
- [118] Shapiro, J. H. *Physical Review A*, **78** 6 061802 (2008) (Cited on page 71.)
- [119] Liu, H., Cheng, J. & Han, S. *Optics communications*, **273** 1 50 (2007) (Cited on page 71.)
- [120] Goodman, J. *Speckle Phenomena in Optics: Theory and Applications* (Roberts & Company) (2010) (Cited on page 72.)
- [121] Erkmen, B. I. & Shapiro, J. H. *Physical Review A*, **77** 4 043809 (2008) (Cited on pages 73, 78, 84, and 86.)
- [122] Erkmen, B. I. & Shapiro, J. H. *Physical Review A*, **79** 2 023833 (2009) (Cited on pages 73 and 83.)
- [123] OSullivan, M. N., Chan, K. W. C. & Boyd, R. W. *Physical Review A*, **82** 5 053803 (2010) (Cited on page 73.)

## REFERENCES

---

- [124] Bennink, R. S., Bentley, S. J., Boyd, R. W. & Howell, J. C. *Physical review letters*, **92** 3 033601 (2004) (Cited on page 73.)
- [125] Brambilla, E., Gatti, A., Bache, M. & Lugiato, L. A. *Physical Review A*, **69** 2 023802 (2004) (Cited on page 73.)
- [126] Gatti, A., Brambilla, E., Bache, M. & Lugiato, L. A. *Physical review letters*, **93** 9 093602 (2004) (Cited on page 73.)
- [127] Brida, G., Chekhova, M., Fornaro, G., Genovese, M., Lopaeva, E. & Berchera, I. R. *Physical Review A*, **83** 6 063807 (2011) (Cited on pages 75, 76, and 84.)
- [128] Rubin, M. H. *Physical Review A*, **54** 6 5349 (1996) (Cited on page 78.)
- [129] Yuen, H. P. & Shapiro, J. H. *Information Theory, IEEE Transactions on*, **24** 6 657 (1978) (Cited on page 78.)
- [130] Shapiro, J. H. & Boyd, R. W. *Quantum Information Processing*, **11** 4 1003 (2012) (Cited on page 86.)
- [131] Audenaert, K., Calsamiglia, J., Muñoz-Tapia, R., Bagan, E., Masanes, L., Acín, A. & Verstraete, F. *Physical review letters*, **98** 16 160501 (2007) (Cited on page 87.)
- [132] Brida, G., Genovese, M. & Berchera, I. R. *Nature Photonics*, **4** 4 227 (2010) (Cited on pages 89 and 93.)
- [133] Brida, G., Caspani, L., Gatti, A., Genovese, M., Meda, A. & Berchera, I. R. *Physical review letters*, **102** 21 213602 (2009) (Cited on pages 89 and 93.)
- [134] Lopaeva, E., Berchera, I. R., Olivares, S., Brida, G., Degiovanni, I. & Genovese, M. *Physica Scripta*, **2014** T160 014026 (2014) (Cited on page 93.)
- [135] Bondani, M., Allevi, A., Zambra, G., Paris, M. G. & Andreoni, A. *Physical Review A*, **76** 1 013833 (2007) (Cited on page 93.)
- [136] Iskhakov, T., Chekhova, M. V. & Leuchs, G. *Physical review letters*, **102** 18 183602 (2009) (Cited on page 93.)

## REFERENCES

---

- [137] Farace, A., De Pasquale, A., Rigovacca, L. & Giovannetti, V. *New Journal of Physics*, **16** 7 073010 (2014) (Cited on page 96.)
- [138] Weedbrook, C., Pirandola, S., Thompson, J., Vedral, V. & Gu, M. *arXiv preprint arXiv:1312.3332* (2013) (Cited on page 96.)
- [139] Gu, M., Chrzanowski, H. M., Assad, S. M., Symul, T., Modi, K., Ralph, T. C., Vedral, V. & Lam, P. K. *Nature Physics*, **8** 9 671 (2012) (Cited on page 96.)
- [140] Barndorff-Nielsen, O. & Gill, R. *Journal of Physics A: Mathematical and General*, **33** 24 4481 (2000) (Cited on page 100.)
- [141] Genoni, M., Paris, M., Adesso, G., Nha, H., Knight, P. & Kim, M. *Physical Review A*, **87** 1 012107 (2013) (Cited on page 102.)
- [142] Macieszczak, K. *arXiv preprint arXiv:1312.1356* (2013) (Cited on page 102.)
- [143] Braunstein, S. L. & Caves, C. M. *Physical Review Letters*, **72** 22 3439 (1994) (Cited on page 103.)
- [144] Paris, M. G. *International Journal of Quantum Information*, **7** supp01 125 (2009) (Cited on page 103.)
- [145] Matsumoto, K. *Journal of Physics A: Mathematical and General*, **35** 13 3111 (2002) (Cited on pages 104, 107, and 113.)
- [146] Hayashi, M. *arXiv preprint arXiv:1011.2546* (2010) (Cited on page 105.)
- [147] Demkowicz-Dobrzański, R., Kołodyński, J. & Guţă, M. *Nature communications*, **3** 1063 (2012) (Cited on page 105.)
- [148] Maccone, L. *arXiv preprint arXiv:1304.7609* (2013) (Cited on page 106.)
- [149] Ballester, M. A. *Physical Review A*, **69** 2 022303 (2004) (Cited on pages 108 and 114.)
- [150] Giovannetti, V., Lloyd, S. & Maccone, L. *Physical review letters*, **96** 1 010401 (2006) (Cited on page 108.)
- [151] Guta, M. Private communication (Cited on page 108.)

## REFERENCES

---

- [152] Guță, M. & Jenčová, A. *Communications in Mathematical Physics*, **276** 2 341 (2007) (Cited on page 108.)
- [153] Vaneph, C., Tufarelli, T. & Genoni, M. G. *Quantum Measurements and Quantum Metrology*, **1** 12 (2013) (Cited on pages 108, 114, and 115.)
- [154] Vidrighin, M. D., Donati, G., Genoni, M. G., Jin, X.-M., Kolthammer, W. S., Kim, M., Datta, A., Barbieri, M. & Walmsley, I. A. *Nature communications*, **5** (2014) (Cited on pages 108, 122, 123, and 133.)
- [155] Yamagata, K., Fujiwara, A., Gill, R. D. et al. *The Annals of Statistics*, **41** 4 2197 (2013) (Cited on page 108.)
- [156] Hayashi, M. *Asymptotic theory of quantum statistical inference*, volume 1 (2005) (Cited on page 111.)
- [157] Fujiwara, A. *Physical Review A*, **65** 1 012316 (2001) (Cited on page 114.)
- [158] Michler, M., Mattle, K., Weinfurter, H. & Zeilinger, A. *Phys. Rev. A*, **53** R1209 (1996) (Cited on page 121.)
- [159] Calsamiglia, J. & Lütkenhaus, N. *Applied Physics B*, **72** 1 67 (2001) (Cited on page 121.)
- [160] Knysh, S. I. & Durkin, G. A. *ArXiv e-prints* (2013). 1307.0470 (Cited on page 122.)
- [161] Huelga, S. F., Macchiavello, C., Pellizzari, T., Ekert, A. K., Plenio, M. & Cirac, J. *Physical Review Letters*, **79** 20 3865 (1997) (Cited on pages 122 and 123.)
- [162] Macieszczak, K. *arXiv preprint arXiv:1403.0955* (2014) (Cited on page 122.)
- [163] Harrow, A. W. *eprint arXiv:quant-ph/0512255* (2005). quant-ph/0512255 (Cited on page 124.)
- [164] Christandl, M. *The Structure of Bipartite Quantum States - Insights from Group Theory and Cryptography*. Ph.D. thesis (2006) (Cited on page 124.)
- [165] Jarzyna, M. & Demkowicz-Dobrzański, R. *New Journal of Physics*, **17** 1 013010 (2015) (Cited on page 124.)

## REFERENCES

---

- [166] Edmonds, A. *Angular Momentum in Quantum Mechanics*. Investigations in Physics Series (Princeton University Press) (1996) (Cited on page 125.)
- [167] Crowley, P. J., Datta, A., Barbieri, M. & Walmsley, I. A. *Physical Review A*, **89** 2 023845 (2014) (Cited on pages 129 and 132.)
- [168] Demkowicz-Dobrzanski, R., Dorner, U., Smith, B., Lundeen, J., Wasilewski, W., Banaszek, K. & Walmsley, I. *Physical Review A*, **80** 1 013825 (2009) (Cited on page 132.)
- [169] Mišta Jr, L., Tatham, R., Girolami, D., Korolkova, N. & Adesso, G. *Physical Review A*, **83** 4 042325 (2011) (Cited on page 135.)
- [170] Aasi, J., Abadie, J., Abbott, B., Abbott, R., Abbott, T., Abernathy, M., Adams, C., Adams, T., Addesso, P., Adhikari, R. et al. *Nature Photonics*, **7** 8 613 (2013) (Cited on page 136.)
- [171] Yuen, H. P. & Shapiro, J. H. *Information Theory, IEEE Transactions on*, **26** 1 78 (1980) (Cited on page 138.)
- [172] Shapiro, J. H. & Sun, K.-X. *JOSA B*, **11** 6 1130 (1994) (Cited on page 138.)